```
  1  GACAGAGTGCAGCCTTTTCAGACTCTGTGACACAGTTCCCCTTTT
 46  GCAAAAATACTTAGCGAGGATCATTACTTTCCAACAGTCGTGTCC
 91  AGAGACCTACTTTGTAACACCGCAGGGAAGTTAATGTACTAGGTC
136  TTGAAAGGTCTTTCTGGAATGTGCAGTAACTTGTAGTTTTCTTCT
181  AGTAGCACTGCTAATTTTTGTGTTATAATTTTTGTAGGTCCATGG

226  GGCCGATGTATGGGAGATGAATGTGGTCCCGGAGGCATCCAAACG
                  MetGlyAspGluCysGlyProGlyGlyIleGlnThr

271  AGGGCTGTGTGGTGTGCTCATGTGGAGGGATGGACTACACTGCAT
     ArgAlaValTrpCysAlaHisValGluGlyTrpThrThrLeuHis

316  ACTAACTGTAAGCAGGCCGAGAGACCCAATAACCAGCAGAATTGT
     ThrAsnCysLysGlnAlaGluArgProAsnAsnGlnGlnAsnCys

361  TTCAAAGTTTGCGATTGGCACAAAGAGTTGTACGACTGGAGACTG
     PheLysValCysAspTrpHisLysGluLeuTyrAspTrpArgLeu

406  GGACCTTGGAATCAGTGTCAGCCCGTGATTTCAAAAAGCCTAGAG
     GlyProTrpAsnGlnCysGlnProValIleSerLysSerLeuGlu

451  AAACCTCTTGAGTGCATTAAGGGGGAAGAAGGTATTCAGGTGAGG
     LysProLeuGluCysIleLysGlyGluGluGlyIleGlnValArg

496  GAGATAGCGTGCATCCAGAAAGACAAAGACATTCCTGCGGAGGAT
     GluIleAlaCysIleGlnLysAspLysAspIleProAlaGluAsp

541  ATCATCTGTGAGTACTTTGAGCCCAAGCCTCTCCTGGAGCAGGCT
     IleIleCysGluTyrPheGluProLysProLeuLeuGluGlnAla

586  TGCCTCATTCCTTGCCAGCAAGATTGCATCGTGTCTGAATTTTCT
     CysLeuIleProCysGlnGlnAspCysIleValSerGluPheSer

631  GCCTGGTCCGAATGCTCCAAGACCTGCGGCAGCGGGCTCCAGCAC
     AlaTrpSerGluCysSerLysThrCysGlySerGlyLeuGlnHis

676  CGGACGCGTCATGTGGTGGCGCCCCCGCAGTTCGGAGGCTCTGGC
     ArgThrArgHisValValAlaProProGlnPheGlyGlySerGly

721  TGTCCAAACCTGACGGAGTTCCAGGTGTGCCAATCCAGTCCATGC
     CysProAsnLeuThrGluPheGlnValCysGlnSerSerProCys

766  GAGGCCGAGGAGCTCAGGTACAGCCTGCATGTGGGGCCCTGGAGC
     GluAlaGluGluLeuArgTyrSerLeuHisValGlyProTrpSer

811  ACCTGCTCAATGCCCCACTCCCGACAAGTAAGACAAGCAAGGAGA
     ThrCysSerMetProHisSerArgGlnValArgGlnAlaArgArg
```

Fig 1

856 CGCGGGAAGAATAAAGAACGGGAAAAGGACCGCAGCAAAGGAGTA
ArgGlyLysAsnLysGluArgGluLysAspArgSerLysGlyVal

901 AAGGATCCAGAAGCCCGCGAGCTTATTAAGAAAAAGAGAAACAGA
LysAspProGluAlaArgGluLeuIleLysLysLysArgAsnArg

946 AACAGGCAGAACAGACAAGAGAACAAATATTGGGACATCCAGATT
AsnArgGlnAsnArgGlnGluAsnLysTyrTrpAspIleGlnIle

991 GGATATCAGACCAGAGAGGTTATGTGCATTAACAAGACGGGGAAA
GlyTyrGlnThrArgGluValMetCysIleAsnLysThrGlyLys

1036 GCTGCTGATTTAAGCTTTTGCCAGCAAGAGAAGCTTCCAATGACC
AlaAlaAspLeuSerPheCysGlnGlnGluLysLeuProMetThr

1081 TTCCAGTCCTGTGTGATCACCAAAGAGTGCCAGGTTTCCGAGTGG
PheGlnSerCysValIleThrLysGluCysGlnValSerGluTrp

1126 TCAGAGTGGAGCCCCTGCTCAAAAACATGCCATGACATGGTGTCC
SerGluTrpSerProCysSerLysThrCysHisAspMetValSer

1171 CCTGCAGGCACTCGTGTAAGGACACGAACCATCAGGCAGTTTCCC
ProAlaGlyThrArgValArgThrArgThrIleArgGlnPhePro

1216 ATTGGCAGTGAAAAGGAGTGTCCAGAATTTGAAGAAAAAGAACCC
IleGlySerGluLysGluCysProGluPheGluGluLysGluPro

1261 TGTTTGTCTCAAGGAGATGGAGTTGTCCCCTGTGCCACGTATGGC
CysLeuSerGlnGlyAspGlyValValProCysAlaThrTyrGly

1306 TGGAGAACTACAGAGTGGACTGAGTGCCGTGTGGACCCTTTGCTC
TrpArgThrThrGluTrpThrGluCysArgValAspProLeuLeu

1351 AGTCAGCAGGACAAGAGGCGCGGCAACCAGACGGCCCTCTGTGGA
SerGlnGlnAspLysArgArgGlyAsnGlnThrAlaLeuCysGly

1396 GGGGGCATCCAGACCCGAGAGGTGTACTGCGTGCAGGCCAACGAA
GlyGlyIleGlnThrArgGluValTyrCysValGlnAlaAsnGlu

1441 AACCTCCTCTCACAATTAAGTACCCACAAGAACAAAGAAGCCTCA
AsnLeuLeuSerGlnLeuSerThrHisLysAsnLysGluAlaSer

1486 AAGCCAATGGACTTAAAATTATGCACTGGACCTATCCCTAATACT
LysProMetAspLeuLysLeuCysThrGlyProIleProAsnThr

1531 ACACAGCTGTGCCACATTCCTTGTCCAACTGAATGTGAAGTTTCA
ThrGlnLeuCysHisIleProCysProThrGluCysGluValSer

Fig 1 (continued)

1576 CCTTGGTCAGCTTGGGGACCTTGTACTTATGAAAACTGTAATGAT
ProTrpSerAlaTrpGlyProCysThrTyrGluAsnCysAsnAsp

1621 CAGCAAGGGAAAAAGGCTTCAAACTGAGGAAGCGGCGCATTACC
GlnGlnGlyLysLysGlyPheLysLeuArgLysArgArgIleThr

1666 AATGAGCCCACTGGAGGCTCTGGGGTAACCGGAAACTGCCCTCAC
AsnGluProThrGlyGlySerGlyValThrGlyAsnCysProHis

1711 TTACTGGAAGCCATTCCCTGTGAAGAGCCTGCCTGTTATGACTGG
LeuLeuGluAlaIleProCysGluGluProAlaCysTyrAspTrp

1756 AAAGCGGTGAGACTGGGAGACTGCGAGCCAGATAACGGAAAGGAG
LysAlaValArgLeuGlyAspCysGluProAspAsnGlyLysGlu

1801 TGTGGTCCAGGCACGCAAGTTCAAGAGGTTGTGTGCATCAACAGT
CysGlyProGlyThrGlnValGlnGluValValCysIleAsnSer

1846 GATGGAGAAGAAGTTGACAGACAGCTGTGCAGAGATGCCATCTTC
AspGlyGluGluValAspArgGlnLeuCysArgAspAlaIlePhe

1891 CCCATCCCTGTGGCCTGTGATGCCCCATGCCCGAAAGACTGTGTG
ProIleProValAlaCysAspAlaProCysProLysAspCysVal

1936 CTCAGCACATGGTCTACGTGGTCCTCCTGCTCACACACCTGCTCA
LeuSerThrTrpSerThrTrpSerSerCysSerHisThrCysSer

1981 GGGAAAACGACAGAAGGGAAACAGATACGAGCACGATCCATTCTG
GlyLysThrThrGluGlyLysGlnIleArgAlaArgSerIleLeu

2026 GCCTATGCGGGTGAAGAAGGTGGAATTCGCTGTCCAAATAGCAGT
AlaTyrAlaGlyGluGluGlyGlyIleArgCysProAsnSerSer

2071 GCTTTGCAAGAAGTACGAAGCTGTAATGAGCATCCTTGCACAGTG
AlaLeuGlnGluValArgSerCysAsnGluHisProCysThrVal

2116 TACCACTGGCAAACTGGTCCCTGGGGCCAGTGCATTGAGGACACC
TyrHisTrpGlnThrGlyProTrpGlyGlnCysIleGluAspThr

2161 TCAGTATCGTCCTTCAACACAACTACGACTTGGAATGGGGAGGCC
SerValSerSerPheAsnThrThrThrThrTrpAsnGlyGluAla

2206 TCCTGCTCTGTCGGCATGCAGACAAGAAAAGTCATCTGTGTGCGA
SerCysSerValGlyMetGlnThrArgLysValIleCysValArg

2251 GTCAATGTGGGCCAAGTGGGACCCAAAAAATGTCCTGAAAGCCTT
ValAsnValGlyGlnValGlyProLysLysCysProGluSerLeu

Fig 1 (continued)

[redacted]

2296 CGACCTGAAACTGTAAGGCCTTGTCTGCTTCCTTGTAAGAAGGAC
ArgProGluThrValArgProCysLeuLeuProCysLysLysAsp

2341 TGTATTGTGACCCCATATAGTGACTGGACATCATGCCCCTCTTCG
CysIleValThrProTyrSerAspTrpThrSerCysProSerSer

2386 TGTAAAGAAGGGGACTCCAGTATCAGGAAGCAGTCTAGGCATCGG
CysLysGluGlyAspSerSerIleArgLysGlnSerArgHisArg

2431 GTCATCATTCAGCTGCCAGCCAACGGGGGCCGAGACTGCACAGAT
ValIleIleGlnLeuProAlaAsnGlyGlyArgAspCysThrAsp

2476 CCCCTCTATGAAGAGAAGGCCTGTGAGGCACCTCAAGCGTGCCAA
ProLeuTyrGluGluLysAlaCysGluAlaProGlnAlaCysGln

2521 AGCTACAGGTGGAAGACTCACAAATGGCGCAGATGCCAATTAGTC
SerTyrArgTrpLysThrHisLysTrpArgArgCysGlnLeuVal

2566 CCTTGGAGCGTGCAACAAGACAGCCCTGGAGCACAGGAAGGCTGT
ProTrpSerValGlnGlnAspSerProGlyAlaGlnGluGlyCys

2611 GGGCCTGGGCGACAGGCAAGAGCCATTACTTGTCGCAAGCAAGAT
GlyProGlyArgGlnAlaArgAlaIleThrCysArgLysGlnAsp

2656 GGAGGACAGGCTGGAATCCATGAGTGCCTACAGTATGCAGGCCCT
GlyGlyGlnAlaGlyIleHisGluCysLeuGlnTyrAlaGlyPro

2701 GTGCCAGCCCTTACCCAGGCCTGCCAGATCCCCTGCCAGGATGAC
ValProAlaLeuThrGlnAlaCysGlnIleProCysGlnAspAsp

2746 TGTCAATTGACCAGCTGGTCCAAGTTTTCTTCATGCAATGGAGAC
CysGlnLeuThrSerTrpSerLysPheSerSerCysAsnGlyAsp

2791 TGTGGTGCAGTTAGGACCAGAAAGCGCACTCTTGTTGGAAAAAGT
CysGlyAlaValArgThrArgLysArgThrLeuValGlyLysSer

2836 AAAAAGAAGGAAAAATGTAAAAATTCCCATTTGTATCCCCTGATT
LysLysLysGluLysCysLysAsnSerHisLeuTyrProLeuIle

2881 GAGACTCAGTATTGTCCTTGTGACAAATATAATGCACAACCTGTG
GluThrGlnTyrCysProCysAspLysTyrAsnAlaGlnProVal

2926 GGGAACTGGTCAGACTGTATTTTACCAGAGGGAAAAGTGGAAGTG
GlyAsnTrpSerAspCysIleLeuProGluGlyLysValGluVal

2971 TTGCTGGGAATGAAAGTACAAGGAGACATCAAGGAATGCGGACAA
LeuLeuGlyMetLysValGlnGlyAspIleLysGluCysGlyGln

Fig 1 (continued)

3016 GGATATCGTTACCAAGCAATGGCATGCTACGATCAAAATGGCAGG
GlyTyrArgTyrGlnAlaMetAlaCysTyrAspGlnAsnGlyArg

3061 CTTGTGGAAACATCTAGATGTAACAGCCATGGTTACATTGAGGAG
LeuValGluThrSerArgCysAsnSerHisGlyTyrIleGluGlu

3106 GCCTGCATCATCCCCTGCCCCTCAGACTGCAAGCTCAGTGAGTGG
AlaCysIleIleProCysProSerAspCysLysLeuSerGluTrp

3151 TCCAACTGGTCGCGCTGCAGCAAGTCCTGTGGGAGTGGTGTGAAG
SerAsnTrpSerArgCysSerLysSerCysGlySerGlyValLys

3196 GTTCGTTCTAAATGGCTGCGTGAAAAACCATATAATGGAGGAAGG
ValArgSerLysTrpLeuArgGluLysProTyrAsnGlyGlyArg

3241 CCTTGCCCCAAACTGGACCATGTCAACCAGGCACAGGTGTATGAG
ProCysProLysLeuAspHisValAsnGlnAlaGlnValTyrGlu

3286 GTTGTCCCATGCCACAGTGACTGCAACCAGTACCTATGGGTCACA
ValValProCysHisSerAspCysAsnGlnTyrLeuTrpValThr

3331 GAGCCCTGGAGCATCTGCAAGGTGACCTTTGTGAATATGCGGGAG
GluProTrpSerIleCysLysValThrPheValAsnMetArgGlu

3376 AACTGTGGAGAGGGCGTGCAAACCCGAAAAGTGAGATGCATGCAG
AsnCysGlyGluGlyValGlnThrArgLysValArgCysMetGln

3421 AATACAGCAGATGGCCCTTCTGAACATGTAGAGGATTACCTCTGT
AsnThrAlaAspGlyProSerGluHisValGluAspTyrLeuCys

3466 GACCCAGAAGAGATGCCCCTGGGCTCTAGAGTGTGCAAATTACCA
AspProGluGluMetProLeuGlySerArgValCysLysLeuPro

3511 TGCCCTGAGGACTGTGTGATATCTGAATGGGGTCCATGGACCCAA
CysProGluAspCysValIleSerGluTrpGlyProTrpThrGln

3556 TGTGTTTTGCCTTGCAATCAAAGCAGTTTCCGGCAAAGGTCAGCT
CysValLeuProCysAsnGlnSerSerPheArgGlnArgSerAla

3601 GATCCCATCAGACAACCAGCTGATGAAGGAAGATCTTGCCCTAAT
AspProIleArgGlnProAlaAspGluGlyArgSerCysProAsn

3646 GCTGTTGAGAAAGAACCCTGTAACCTGAACAAAAACTGCTACCAC
AlaValGluLysGluProCysAsnLeuAsnLysAsnCysTyrHis

3691 TATGATTATAATGTAACAGACTGGAGTACATGTCAGCTGAGTGAG
TyrAspTyrAsnValThrAspTrpSerThrCysGlnLeuSerGlu

Fig 1 (continued)

3736 AAGGCAGTTTGTGGAAATGGAATAAAAACAAGGATGTTGGATTGT
LysAlaValCysGlyAsnGlyIleLysThrArgMetLeuAspCys

3781 GTTCGAAGTGATGGCAAGTCAGTTGACCTGAAATATTGTGAAGCG
ValArgSerAspGlyLysSerValAspLeuLysTyrCysGluAla

3826 CTTGGCTTGGAGAAGAACTGGCAGATGAACACGTCCTGCATGGTG
LeuGlyLeuGluLysAsnTrpGlnMetAsnThrSerCysMetVal

3871 GAATGCCCTGTGAACTGTCAGCTTTCTGATTGGTCTCCTTGGTCA
GluCysProValAsnCysGlnLeuSerAspTrpSerProTrpSer

3916 GAATGTTCTCAAACATGTGGCCTCACAGGAAAAATGATCCGAAGA
GluCysSerGlnThrCysGlyLeuThrGlyLysMetIleArgArg

3961 CGAACAGTGACCCAGCCCTTTCAAGGTGATGGAAGACCATGCCCT
ArgThrValThrGlnProPheGlnGlyAspGlyArgProCysPro

4006 TCCCTGATGGACCAGTCCAAACCCTGCCCAGTGAAGCCTTGTTAT
SerLeuMetAspGlnSerLysProCysProValLysProCysTyr

4051 CGGTGGCAATATGGCCAGTGGTCTCCATGCCAAGTGCAGGAGGCC
ArgTrpGlnTyrGlyGlnTrpSerProCysGlnValGlnGluAla

4096 CAGTGTGGAGAAGGGACCAGAACAAGGAACATTTCTTGTGTAGTA
GlnCysGlyGluGlyThrArgThrArgAsnIleSerCysValVal

4141 AGTGATGGGTCAGCTGATGATTTCAGCAAAGTGGTGGATGAGGAA
SerAspGlySerAlaAspAspPheSerLysValValAspGluGlu

4186 TTCTGTGCTGACATTGAACTCATTATAGATGGTAATAAAAATATG
PheCysAlaAspIleGluLeuIleIleAspGlyAsnLysAsnMet

4231 GTTCTGGAGGAATCCTGCAGCCAGCCTTGCCCAGGTGACTGTTAT
ValLeuGluGluSerCysSerGlnProCysProGlyAspCysTyr

4276 TTGAAGGACTGGTCTTCCTGGAGCCTGTGTCAGCTGACCTGTGTG
LeuLysAspTrpSerSerTrpSerLeuCysGlnLeuThrCysVal

4321 AATGGTGAGGATCTAGGCTTTGGTGGAATACAGGTCAGATCCAGA
AsnGlyGluAspLeuGlyPheGlyGlyIleGlnValArgSerArg

4366 CCGGTGATTATACAAGAACTAGAGAATCAGCATCTGTGCCCAGAG
ProValIleIleGlnGluLeuGluAsnGlnHisLeuCysProGlu

4411 CAGATGTTAGAAACAAAATCATGTTATGATGGACAGTGCTATGAA
GlnMetLeuGluThrLysSerCysTyrAspGlyGlnCysTyrGlu

4456 TATAAATGGATGGCCAGTGCTTGGAAGGGCTCTTCCCGAACAGTG
TyrLysTrpMetAlaSerAlaTrpLysGlySerSerArgThrVal

Fig 1 (continued)

4501 TGGTGTCAAAGGTCAGATGGTATAAATGTAACAGGGGGCTGCTTG
TrpCysGlnArgSerAspGlyIleAsnValThrGlyGlyCysLeu

4546 GTGATGAGCCAGCCTGATGCCGACAGGTCTTGTAACCCACCGTGT
ValMetSerGlnProAspAlaAspArgSerCysAsnProProCys

4591 AGTCAACCCCACTCGTACTGTAGCGAGACAAAAACATGCCATTGT
SerGlnProHisSerTyrCysSerGluThrLysThrCysHisCys

4636 GAAGAAGGGTACACTGAAGTCATGTCTTCTAACAGCACCCTTGAG
GluGluGlyTyrThrGluValMetSerSerAsnSerThrLeuGlu

4681 CAATGCACACTTATCCCCGTGGTGGTATTACCCACCATGGAGGAC
GlnCysThrLeuIleProValValValLeuProThrMetGluAsp

4726 AAAAGAGGAGATGTGAAAACCAGTCGGGCTGTACATCCAACCCAA
LysArgGlyAspValLysThrSerArgAlaValHisProThrGln

4771 CCCTCCAGTAACCCAGCAGGACGGGGAAGGACCTGGTTTCTACAG
ProSerSerAsnProAlaGlyArgGlyArgThrTrpPheLeuGln

4816 CCATTTGGGCCAGATGGGAGACTAAAGACCTGGGTTTACGGTGTA
ProPheGlyProAspGlyArgLeuLysThrTrpValTyrGlyVal

4861 GCAGCTGGGGCATTTGTGTTACTCATCTTTATTGTCTCCATGATT
AlaAlaGlyAlaPheValLeuLeuIlePheIleValSerMetIle

4906 TATCTAGCTTGCAAAAAGCCAAAGAAACCCCAAAGAAGGCAAAAC
TyrLeuAlaCysLysLysProLysLysProGlnArgArgGlnAsn

4951 AACCGACTGAAACCTTTAACCTTAGCCTATGATGGAGATGCCGAC
AsnArgLeuLysProLeuThrLeuAlaTyrAspGlyAspAlaAsp

4996 ATGTAACATATAACTTTTCCTGGCAACAACCAGTTTCGGCTTTCT
Met

Fig 1 (continued)

5041 GACTTCATAGATGTCCAGAGGCCACAACAAATGTATCCAAACTGT
5086 GTGGATTAAAATATATTTTAATTTTTAAAAATGGCATCATAAAGA
5131 CAAGAGTGAAAATCATACTGCCACTGGAGATATTTAAGACAGTAC
5176 CACTTATATACAGACCATCAACCGTGAGAATTATAGGAGATTTAG
5221 CTGAATACATGCTGCATTCTGAAAGTTTTATGTCATCTTTTCTGA
5266 AATCTACCGACTGAAAAACCACTTTCATCTCTAAAAAATAATGGT
5311 GGAATTGGCCAGTTAGGATGCCTGATACAAGACCGTCTGCAGTGT
5356 TAATCCATAAAACTTCCTAGCATGAAGAGTTTCTACCAAGATCTC
5401 CACAATACTATGGTCAAATTAACATGTGTACTCAGTTGAATGACA
5446 CACATTATGTCAGATTATGTACTTGCTAATAAGCAATTTTAACAA
5491 TGCATAACAAATAAACTCTAAGCTAAGCAGAAAATCCACTGAATA
5536 AATTCAGCATCTTGGTGGTCGATGGTAGATTTTATTGACCTGCAT
5581 TTCAGAGACAAAGCCTCTTTTTTAAGACTTCTTGTCTCTCTCCAA
5626 AGTAAGAATGCTGGACAAGTACTAGTGTCTTAGAAGAACGAGTCC
5671 TCAAGTTCAGTATTTTATAGTGGTAATTGTCTGGAAAACTAATTT
5716 ACTTGTGTTAATACAATACGTTTCTACTTTCCCTGATTTTCAAAC
5761 TGGTTGCCTGCATCTTTTTTGCTATATGGAAGGCACATTTTTGCA
5806 CTATATTAGTGCAGCACGATAGGCGCTTAACCAGTATTGCCATAG
5851 AAACTGCCTCTTTTCATGTGGGATGAAGACATCTGTGCCAAGAGT
5896 GGCATGAAGACATTTGCAAGTTCTTGTATCCTGAAGAGAGTAAAG
5941 TTCAGTTTGGATGGCAGCAAGATGAAATCAGCTATTACACCTGCT
5986 GTACACACACTTCCTCATCACTGCAGCCATTGTGAAATTGACAAC
6031 ATGGCGGTAATTTAAGTGTTGAAGTCCCTAACCCCTTAACCCTCT
6076 AAAAGGTGGATTCCTCTAGTTGGTTTGTAATTGTTCTTTGAAGGC
6121 TGTTTATGACTAGATTTTTATATTTGTTATCTTTGTTAAGAAAAA
6166 AAAAAGAAAAAGGAACTGGATGTCTTTTTAATTTTGAGCAGATGG
6211 AGAAAATAAATAATGTATCAATGACCTTTGTAACTAAAGGAAAAA
6256 AAAAAAAAATGTGGATTTTCCTTTCTCTCTGATTTCCCAGTTTCA
6301 GATTGAATGTCTGTCTTGCAGGCAGTTATTTCAAAATCCATAGTC
6346 TTTNGCCTTTCTCACTGGCAAAATTTGA

Fig 1 (continued)

1 CACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCCCCTTTGGAG
46 CCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAGTGATGGTA
91 AGTGCTGGCCCAGACTGAAGCTCGGAGAGGCACTCTGCTTGCCCA
136 GCGTCACAGTCTTAGCTCCCAACTGTCCTGGCTTCCAGTCTCCCT
181 TGCTTCCCAGATCCCAGACTCTAGCCCCAGCCCCGTCTCTTTCAC
226 CAGCTCCTGGGACCCTACGCAATCTGCGCCTGCGTCTCATCAGTC
271 GCCCCACATGTAACTGTATCTACAACCAGCTGCACCAGCGACACC
316 TGTCCAACCCGGCCCGGCCTGGGATGCTATGTGGGGCCCCCAGC
361 CTGGGGTGCAGGGCCCCTGTCAGGTCTGATAGGGAGAAGAGAAGG
406 AGCAGAAGGGGAGGGGCCTAACCCTGGGCTGGGGGTTGGACTCAC
451 AGGACTGGGGGAAAGAGCTGCAATCAGAGGGTGTCTGCCATAGCT
496 GGGCTCAGGCATCTGTCCTTGGCTTTGTTGCCTGGCTCCAGGGAG
541 ATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGACACTGGG
586 TTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGG
631 ACGCTCCTGTGCTGCTGACCAACACAGCTGCTCACAGTTCCTGGC
676 TGCAGGCTCGAGTTCAGGGGGCAGCTTTCCTGGCCCAGAGCCCAG

721 AGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGAT
MetSerAspGluAspSerCysValAlaCysGlyS

766 CCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCCTCCCCATGGC
erLeuArgThrAlaGlyProGlnAlaGlyAlaProSerProTrpP

811 CCTGGGAGGCCAGGCTGATGCACCAGGGACAGCTGGCCTGTGGCG
roTrpGluAlaArgLeuMetHisGlnGlyGlnLeuAlaCysGlyG

856 GAGCCCTGGTGTCAGAGGAGGCGGTGCTAACTGCTGCCCACTGCT
lyAlaLeuValSerGluGluAlaValLeuThrAlaAlaHisCysP

901 TCAATGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGA
heAsnGlyArgGlnAlaProGluGluTrpSerValGlyLeuGlyT

946 CCAGACCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAG
hrArgProGluGluTrpGlyLeuLysGlnLeuIleLeuHisGlyA

991 CCTACACCCACCCTGAGGGGGGCTACGACATGGCCCTCCTGCTGC
laTyrThrHisProGluGlyGlyTyrAspMetAlaLeuLeuLeuL

1036 TGGCTCAGCCTGTGACACTGGGAGCCAGCCTGCGGGCCCTCTGCC
euAlaGlnProValThrLeuGlyAlaSerLeuArgAlaLeuCysL

1081 TGCCCTATTTTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGG
euProTyrPheAspHisHisLeuProAspGlyGluArgGlyTrpV

1126 TTCTGGGACGGGCCCGCCCAGGAGCAGGCATCAGCTCCCTCCAGA
alLeuGlyArgAlaArgProGlyAlaGlyIleSerSerLeuGlnT

1171 CAGTGCCCGTGACCCTCCTGGGGCCTAGGGCCTGCAGCCGGCTGC
hrValProValThrLeuLeuGlyProArgAlaCysSerArgLeuH

Fig. 2

```
1216   ATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
       isAlaAlaProGlyGlyAspGlySerProIleLeuProGlyMetV

1261   TGTGTACCAGTGCTGTGGGTGAGCTGCCCAGCTGTGAGGGCCTGT
       alCysThrSerAlaValGlyGluLeuProSerCysGluGlyLeuS

1306   CTGGGGCACCACTGGTGCATGAGGTGAGGGGCACATGGTTCCTGG
       erGlyAlaProLeuValHisGluValArgGlyThrTrpPheLeuA

1351   CCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGGCCCCGCCAGGC
       laGlyLeuHisSerPheGlyAspAlaCysGlnGlyProAlaArgP

1396   CGGCGGTCTTCACCGCGCTCCCTGCCTATGAGGACTGGGTCAGCA
       roAlaValPheThrAlaLeuProAlaTyrGluAspTrpValSerS

1441   GTTTGGACTGGCAGGTCTACTTCGCCGAGGAACCAGAGCCCGAGG
       erLeuAspTrpGlnValTyrPheAlaGluGluProGluProGluA

1486   CTGAGCCTGGAAGCTGCCTGGCCAACATAAGCCAACCAACCAGCT
       laGluProGlySerCysLeuAlaAsnIleSerGlnProThrSerC

1531   GCTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGCAGGCA
       ys

1576   GGCAAATGGCATTACTGCCCCTGTCCTCCCCACCCTGTCATGTGT
1621   GATTCCAGGCACCAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGA
1666   AGGAACCTGCCTGGGGCCACAGGTGCCCCCTCCCCACCCTGCAGG
1711   ACAGGGGTGTCTGTGGACACTCCCACACCCAACTCTGCTACCAAG
1756   CAGGCGTCTCAGCTTTCCTCCTCCTTTACCCTTTCAGATACAATC
1801   ACGCCAGCCCCGTTGTTTTGAAAATTTCTTTTTTTGGGGGGCAGC
1846   AGTTTTCCTTTTTTTAAACTTAAATAAATTGTTACAAAATAGACT
1891   TTAG
```

Fig. 2 (continued)

```
  1  GCGGATCCTCACACGACTGTGATCCGATTCTTTCCAGCGGCTTCT
 46  GCAACCAAGCGGGTCTTACCCCCGGTCCTCCGCGTCTCCAGTCCT
 91  CGCACCTGGAACCCCAACGTCCCCGAGAGTCCCCGAATCCCCGCT

136  CCCAGGCTACCTAAGAGGATGAGCGGTGCTCCGACGGCCGGGGCA
                                 MetSerGlyAlaProThrAlaGlyAla

181  GCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCTCAG
     AlaLeuMetLeuCysAlaAlaThrAlaValLeuLeuSerAlaGln

226  GGCGGACCCGTGCAGTCCAAGTCGCCGCGCTTTGCGTCCTGGGAC
     GlyGlyProValGlnSerLysSerProArgPheAlaSerTrpAsp

271  GAGATGAATGTCCTGGCGCACGGACTCCTGCAGCTCGGCCAGGGG
     GluMetAsnValLeuAlaHisGlyLeuLeuGlnLeuGlyGlnGly

316  TGCGCGAACACCGGAGCGCACCCGCAGTCAGCTGAGCGCGCTGGA
     CysAlaAsnThrGlyAlaHisProGlnSerAlaGluArgAlaGly

361  GCGCGCCTGAGCGCGTGCGGGTCCGCCTGTCAGGGAACCGAGGGG
     AlaArgLeuSerAlaCysGlySerAlaCysGlnGlyThrGluGly

406  TCCACCGACCTCCCGTTAGCCCCTGAGAGCCGGGTGGACCCTGAG
     SerThrAspLeuProLeuAlaProGluSerArgValAspProGlu

451  GTCCTTCACAGCCTGCAGACACAACTCAAGGCTCAGAACAGCAGG
     ValLeuHisSerLeuGlnThrGlnLeuLysAlaGlnAsnSerArg

496  ATCCAGCAACTCTTCCACAAGGTGGCCCAGCAGCAGCGGCACCTG
     IleGlnGlnLeuPheHisLysValAlaGlnGlnGlnArgHisLeu

541  GAGAAGCAGCACCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGC
     GluLysGlnHisLeuArgIleGlnHisLeuGlnSerGlnPheGly

586  CTCCTGGACCACAAGCACCTAGACCATGAGGTGGCCAAGCCTGCC
     LeuLeuAspHisLysHisLeuAspHisGluValAlaLysProAla

631  CGAAGAAAGAGGCTGCCCGAGATGGCCCAGCCAGTTGACCCGGCT
     ArgArgLysArgLeuProGluMetAlaGlnProValAspProAla

676  CACAATGTCAGCCGCCTGCACCGGCTGCCCAGGGATTGCCAGGAG
     HisAsnValSerArgLeuHisArgLeuProArgAspCysGlnGlu

721  CTGTTCCAGGTTGGGGAGAGGCAGAGTGGACTATTTGAAATCCAG
     LeuPheGlnValGlyGluArgGlnSerGlyLeuPheGluIleGln

766  CCTCAGGGGTCTCCGCCATTTTTGGTGAACTGCAAGATGACCTCA
     ProGlnGlySerProProPheLeuValAsnCysLysMetThrSer
```

Fig. 3

```
811   GATGGAGGCTGGACAGTAATTCAGAGGCGCCACGATGGCTCAGTG
      AspGlyGlyTrpThrValIleGlnArgArgHisAspGlySerVal

856   GACTTCAACCGGCCCTGGGAAGCCTACAAGGCGGGGTTTGGGGAT
      AspPheAsnArgProTrpGluAlaTyrLysAlaGlyPheGlyAsp

901   CCCCACGGCGAGTTCTGGCTGGGTCTGGAGAAGGTGCATAGCATG
      ProHisGlyGluPheTrpLeuGlyLeuGluLysValHisSerMet

946   ATGGGGGACCGCAACAGCCGCCTGGCCGTGCAGCTGCGGGACTGG
      MetGlyAspArgAsnSerArgLeuAlaValGlnLeuArgAspTrp

991   GATGGCAACGCCGAGTTGCTGCAGTTCTCCGTGCACCTGGGTGGC
      AspGlyAsnAlaGluLeuLeuGlnPheSerValHisLeuGlyGly

1036  GAGGACACGGCCTATAGCCTGCAGCTCACTGCACCCGTGGCCGGC
      GluAspThrAlaTyrSerLeuGlnLeuThrAlaProValAlaGly

1081  CAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGTACCC
      GlnLeuGlyAlaThrThrValProProSerGlyLeuSerValPro

1126  TTCTCCACTTGGGACCAGGATCACGACCTCCGCAGGGACAAGAAC
      PheSerThrTrpAspGlnAspHisAspLeuArgArgAspLysAsn

1171  TGCGCCAAGAGCCTCTCTGGAGGCTGGTGGTTTGGCACCTGCAGC
      CysAlaLysSerLeuSerGlyGlyTrpTrpPheGlyThrCysSer

1216  CATTCCAACCTCAACGGCCAGTACTTCCGCTCCATCCCACAGCAG
      HisSerAsnLeuAsnGlyGlnTyrPheArgSerIleProGlnGln

1261  CGGCAGAAGCTTAAGAAGGGAATCTTCTGGAAGACCTGGCGGGGC
      ArgGlnLysLeuLysLysGlyIlePheTrpLysThrTrpArgGly

1306  CGCTACTACCCGCTGCAGGCCACCACCATGTTGATCCAGCCCATG
      ArgTyrTyrProLeuGlnAlaThrThrMetLeuIleGlnProMet

1351  GCAGCAGAGGCAGCCTCCTAGCGTCCTGGCTGGGCCTGGTCCCAG
      AlaAlaGluAlaAlaSer

1396  GCCCACGAAAGACGGTGACTCTTGGCTCTGCCCGAGGATGTGGCC
1441  GTTCCCTGCCTGGGCAGGGGCTCCAAGGAGGGGCCATCTGGAAAC
1486  TTGTGGACAGAGAAGAAGACCACGACTGGAGAAGCCCCCTTTCTG
1531  AGTGCAGGGGGGCTGCATGCGTTGCCTCCTGAGATCGAGGCTGCA
1576  GGATATGCTCAGACTCTAGAGGCGTGGACCAAGGGGCATGGAGCT
1621  TCACTCCTTGCTGGCCAGGGAGTTGGGGACTCAGAGGGACCACTT
1666  GGGGCCAGCCAGACTGGCCTCAATGGCGGACTCAGTCACATTGAC
1711  TGACGGGGACCAGGGCTTGTGTGGGTCGAGAGCGCCCTCATGGTG
1756  CTGGTGCTGTTGTGTGTAGGTCCCCTGGGGACACAAGCAGGCGCC
1801  AATGGTATCTGGGCGGAGCTCACAGAGTTCTTGGAATAAAAGCAA
1846  CCTCAGAACA
```

Fig. 3 (continued)

```
  1  GGTAGCCGACGCGCCGGCCGGCGCGTGACCTTGCCCCTCTTGCTC

 46  GCCTTGAAAATGGAAAAGATGCTCGCAGGCTGCTTTCTGCTGATC
                 MetGluLysMetLeuAlaGlyCysPheLeuLeuIle

 91  CTCGGACAGATCGTCCTCCTCCCTGCCGAGGCCAGGGAGCGGTCA
     LeuGlyGlnIleValLeuLeuProAlaGluAlaArgGluArgSer

136  CGTGGGAGGTCCATCTCTAGGGGCAGACACGCTCGGACCCACCCG
     ArgGlyArgSerIleSerArgGlyArgHisAlaArgThrHisPro

181  CAGACGGCCCTTCTGGAGAGTTCCTGTGAGAACAAGCGGGCAGAC
     GlnThrAlaLeuLeuGluSerSerCysGluAsnLysArgAlaAsp

226  CTGGTTTTCATCATTGACAGCTCTCGCAGTGTCAACACCCATGAC
     LeuValPheIleIleAspSerSerArgSerValAsnThrHisAsp

271  TATGCAAAGGTCAAGGAGTTCATCGTGGACATCTTGCAATTCTTG
     TyrAlaLysValLysGluPheIleValAspIleLeuGlnPheLeu

316  GACATTGGTCCTGATGTCACCCGAGTGGGCCTGCTCCAATATGGC
     AspIleGlyProAspValThrArgValGlyLeuLeuGlnTyrGly

361  AGCACTGTCAAGAATGAGTTCTCCCTCAAGACCTTCAAGAGGAAG
     SerThrValLysAsnGluPheSerLeuLysThrPheLysArgLys

406  TCCGAGGTGGAGCGTGCTGTCAAGAGGATGCGGCATCTGTCCACG
     SerGluValGluArgAlaValLysArgMetArgHisLeuSerThr

451  GGCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCA
     GlyThrMetThrGlyLeuAlaIleGlnTyrAlaLeuAsnIleAla

496  TTCTCAGAAGCAGAGGGGGCCCGGCCCCTGAGGGAGAATGTGCCA
     PheSerGluAlaGluGlyAlaArgProLeuArgGluAsnValPro

541  CGGGTCATAATGATCGTGACGGATGGGAGACCTCAGGACTCCGTG
     ArgValIleMetIleValThrAspGlyArgProGlnAspSerVal

586  GCCGAGGTGGCTGCTAAGGCACGGGACACGGGCATCCTAATCTTT
     AlaGluValAlaAlaLysAlaArgAspThrGlyIleLeuIlePhe

631  GCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATT
     AlaIleGlyValGlyGlnValAspPheAsnThrLeuLysSerIle

676  GGGAGTGAGCCCCATGAGGACCATGTCTTCCTTGTGGCCAATTTC
     GlySerGluProHisGluAspHisValPheLeuValAlaAsnPhe

721  AGCCAGATTGAGACGCTGACCTCCGTGTTCCAGAAGAAGTTGTGC
     SerGlnIleGluThrLeuThrSerValPheGlnLysLysLeuCys
```

Fig. 4

```
766   ACGGCCCACATGTGCAGCACCCTGGAGCATAACTGTGCCCACTTC
      ThrAlaHisMetCysSerThrLeuGluHisAsnCysAlaHisPhe

811   TGCATCAACATCCCTGGCTCATACGTCTGCAGGTGCAAACAAGGC
      CysIleAsnIleProGlySerTyrValCysArgCysLysGlnGly

856   TACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTG
      TyrIleLeuAsnSerAspGlnThrThrCysArgIleGlnAspLeu

901   TGTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTG
      CysAlaMetGluAspHisAsnCysGluGlnLeuCysValAsnVal

946   CCGGGCTCCTTCGTCTGCGAGTGCTACAGTGGCTACGCCCTGGCT
      ProGlySerPheValCysGluCysTyrSerGlyTyrAlaLeuAla

991   GAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCTCAGAA
      GluAspGlyLysArgCysValAlaValAspTyrCysAlaSerGlu

1036  AACCACGGATGTGAACATGAGTGTGTAAATGCTGATGGCTCCTAC
      AsnHisGlyCysGluHisGluCysValAsnAlaAspGlySerTyr

1081  CTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAA
      LeuCysGlnCysHisGluGlyPheAlaLeuAsnProAspGluLys

1126  ACGTGCACAAAGATAGACTACTGTGCCTCATCTAATCATGGATGT
      ThrCysThrLysIleAspTyrCysAlaSerSerAsnHisGlyCys

1171  CAGTACGAGTGTGTTAACACAGATGATTCCTATTCCTGCCACTGC
      GlnTyrGluCysValAsnThrAspAspSerTyrSerCysHisCys

1216  CTGAAAGGCTTTACCCTGAATCCAGATAAGAAAACCTGCAGAAGG
      LeuLysGlyPheThrLeuAsnProAspLysLysThrCysArgArg

1261  ATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGC
      IleAsnTyrCysAlaLeuAsnLysProGlyCysGluHisGluCys

1306  GTCAACATGGAGGAGAGCTACTACTGCCGCTGCCACCGTGGCTAC
      ValAsnMetGluGluSerTyrTyrCysArgCysHisArgGlyTyr

1351  ACTCTGGACCCCAATGGCAAACCCTGCAGCCGAGTGGACCACTGT
      ThrLeuAspProAsnGlyLysProCysSerArgValAspHisCys

1396  GCACAGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAG
      AlaGlnGlnAspHisGlyCysGluGlnLeuCysLeuAsnThrGlu

1441  GATTCCTTCGTCTGCCAGTGCTCAGAAGGCTTCCTCATCAACGAG
      AspSerPheValCysGlnCysSerGluGlyPheLeuIleAsnGlu
```

Fig. 4 (continued)

1486 GACCTCAAGACCTGCTCCCGGGTGGATTACTGCCTGCTGAGTGAC
AspLeuLysThrCysSerArgValAspTyrCysLeuLeuSerAsp

1531 CATGGTTGTGAATACTCCTGTGTCAACATGGACAGATCCTTTGCC
HisGlyCysGluTyrSerCysValAsnMetAspArgSerPheAla

1576 TGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAGACG
CysGlnCysProGluGlyHisValLeuArgSerAspGlyLysThr

1621 TGTGCAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAA
CysAlaLysLeuAspSerCysAlaLeuGlyAspHisGlyCysGlu

1666 CATTCGTGTGTAAGCAGTGAAGATTCGTTTGTGTGCCAGTGCTTT
HisSerCysValSerSerGluAspSerPheValCysGlnCysPhe

1711 GAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAA
GluGlyTyrIleLeuArgGluAspGlyLysThrCysArgArgLys

1756 GATGTCTGCCAAGCTATAGACCATGGCTGTGAACACATTTGTGTG
AspValCysGlnAlaIleAspHisGlyCysGluHisIleCysVal

1801 AACAGTGACGACTCATACACGTGCGAGTGCTTGGAGGGATTCCGG
AsnSerAspAspSerTyrThrCysGluCysLeuGluGlyPheArg

1846 CTCACTGAGGATGGGAAACGCTGCCGAATTTCCTCAGGGAAGGAT
LeuThrGluAspGlyLysArgCysArgIleSerSerGlyLysAsp

1891 GTCTGCAAATCAACCCACCATGGCTGCGAACACATTTGTGTTAAT
ValCysLysSerThrHisHisGlyCysGluHisIleCysValAsn

1936 AATGGGAATTCCTACATCTGCAAATGCTCAGAGGGATTTGTTCTA
AsnGlyAsnSerTyrIleCysLysCysSerGluGlyPheValLeu

1981 GCTGAGGACGGAAGACGGTGCAAGAAATGCACTGAAGGCCCAATT
AlaGluAspGlyArgArgCysLysLysCysThrGluGlyProIle

2026 GACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAG
AspLeuValPheValIleAspGlySerLysSerLeuGlyGluGlu

2071 AATTTTGAGGTCGTGAAGCAGTTTGTCACTGGAATTATAGATTCC
AsnPheGluValValLysGlnPheValThrGlyIleIleAspSer

2116 TTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGTAT
LeuThrIleSerProLysAlaAlaArgValGlyLeuLeuGlnTyr

2161 TCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCA
SerThrGlnValHisThrGluPheThrLeuArgAsnPheAsnSer

2206 GCCAAAGACATGAAAAAAGCCGTGGCCCACATGAAATACATGGGA
AlaLysAspMetLysLysAlaValAlaHisMetLysTyrMetGly

Fig. 4 (continued)

2251 AAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGA
LysGlySerMetThrGlyLeuAlaLeuLysHisMetPheGluArg

2296 AGTTTTACCCAAGGAGAAGGGGCCAGGCCCCTTTTCCACAAGGGT
SerPheThrGlnGlyGluGlyAlaArgProLeuPheHisLysGly

2341 GCCCAGAGCAGCCATTGTGTTCACCGACGGACGGGCTCAGGATGA
AlaGlnSerSerHisCysValHisArgArgThrGlySerGly

2386 CGTCTCCGAGTGGGCCAGTAAAGCCAAGGCCAATGGTATCACTAT
2431 GTATGCTGTTGGGGTAGGAAAAGCCATTGAGGAGGAACTACAAGA
2476 GATTGCCTCTGAGCCCACAAACAAGCATCTCTTCTATGCCGAAGA
2521 CTTCAGCACAATGGATGAGATAAGTGAAAAACTCAAGAAAGGCAT
2566 CTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAGC
2611 AGGGGAACTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCC
2656 AGTCACCATAAATATCCAAGACCTACTTTCCTGTTCTAATTTTGC
2701 AGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTC
2746 TACACAAAAGCTTTCCCATTCAACAAAACCTTCAGGAAGCCCTTT
2791 GGAAGAAAAACACGATCAATGCAAATGTGAAAACCTTATAATGTT
2836 CCAGAACCTTGCAAACGAAGAAGTAAGAAAATTTACACAGCGCTT
2881 AGAAGAAATGACACAGAGAATGGAAGCCCTGGAAAATCGCCTGAG
2926 ATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTA
2971 TCACGGATTACAATGAACGCAGTGCAGAGCCCCAAAGCTCAGGCT
3016 ATTGTTAAATC

Fig. 4 (continued)

```
1    GGTAGCCGACGCGCCGGCCGGCGCGTGACCTTGCCCCTCTTGCTC

46   GCCTTGAAAATGGAAAAGATGCTCGCAGGCTGCTTTCTGCTGATC
               MetGluLysMetLeuAlaGlyCysPheLeuLeuIle

91   CTCGGACAGATCGTCCTCCTCCCCTGCGAGGCCAGGGAGCGGTCA
     LeuGlyGlnIleValLeuLeuProCysGluAlaArgGluArgSer

136  CGTGGGAGGTCCATCTCTAGGGGCAGACACGCTCGGACCCACCCG
     ArgGlyArgSerIleSerArgGlyArgHisAlaArgThrHisPro

181  CAGACGGCCCTTCTGGAGAGTTCCTGTGAGAACAAGCGGGCAGAC
     GlnThrAlaLeuLeuGluSerSerCysGluAsnLysArgAlaAsp

226  CTGGTTTTCATCATTGACAGCTCTCGCAGTGTCAACACCCATGAC
     LeuValPheIleIleAspSerSerArgSerValAsnThrHisAsp

271  TATGCAAAGGTCAAGGAGTTCATCGTGGACATCTTGCAATTCTTG
     TyrAlaLysValLysGluPheIleValAspIleLeuGlnPheLeu

316  GACATTGGTCCTGATGTCACCCGAGTGGGCCTGCTCCAATATGGC
     AspIleGlyProAspValThrArgValGlyLeuLeuGlnTyrGly

361  AGCACTGTCAAGAATGAGTTCTCCCTCAAGACCTTCAAGAGGAAG
     SerThrValLysAsnGluPheSerLeuLysThrPheLysArgLys

406  TCCGAGGTGGAGCGTGCTGTCAAGAGGATGCGGCATCTGTCCACG
     SerGluValGluArgAlaValLysArgMetArgHisLeuSerThr

451  GGCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCA
     GlyThrMetThrGlyLeuAlaIleGlnTyrAlaLeuAsnIleAla

496  TTCTCAGAAGCAGAGGGGGCCCGGCCCCTGAGGGAGAATGTGCCA
     PheSerGluAlaGluGlyAlaArgProLeuArgGluAsnValPro

541  CGGGTCATAATGATCGTGACGGATGGGAGACCTCAGGACTCCGTG
     ArgValIleMetIleValThrAspGlyArgProGlnAspSerVal

586  GCCGAGGTGGCTGCTAAGGCACGGGACACGGGCATCCTAATCTTT
     AlaGluValAlaAlaLysAlaArgAspThrGlyIleLeuIlePhe

631  GCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATT
     AlaIleGlyValGlyGlnValAspPheAsnThrLeuLysSerIle

676  GGGAGTGAGCCCCATGAGGACCATGTCTTCCTTGTGGCCAATTTC
     GlySerGluProHisGluAspHisValPheLeuValAlaAsnPhe

721  AGCCAGATTGAGACGCTGACCTCCGTGTTCCAGAAGAAGTTGTGC
     SerGlnIleGluThrLeuThrSerValPheGlnLysLysLeuCys
```

Fig. 5

766 ACGGCCCACATGTGCAGCACCCTGGAGCATAACTGTGCCCACTTC
ThrAlaHisMetCysSerThrLeuGluHisAsnCysAlaHisPhe

811 TGCATCAACATCCCTGGCTCATACGTCTGCAGGTGCAAACAAGGC
CysIleAsnIleProGlySerTyrValCysArgCysLysGlnGly

856 TACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTG
TyrIleLeuAsnSerAspGlnThrThrCysArgIleGlnAspLeu

901 TGTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTG
CysAlaMetGluAspHisAsnCysGluGlnLeuCysValAsnVal

946 CCGGGCTCCTTCGTCTGCGAGTGCTACAGTGGCTACGCCCTGGCT
ProGlySerPheValCysGluCysTyrSerGlyTyrAlaLeuAla

991 GAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCTCAGAA
GluAspGlyLysArgCysValAlaValAspTyrCysAlaSerGlu

1036 AACCACGGATGTGAACATGAGTGTGTAAATGCTGATGGCTCCTAC
AsnHisGlyCysGluHisGluCysValAsnAlaAspGlySerTyr

1081 CTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCAGATGAAAAA
LeuCysGlnCysHisGluGlyPheAlaLeuAsnProAspGluLys

1126 ACGTGCACAAAGATAGACTACTGTGCCTCATCTAATCATGGATGT
ThrCysThrLysIleAspTyrCysAlaSerSerAsnHisGlyCys

1171 CAGTACGAGTGTGTTAACACAGATGATTCCTATTCCTGCCACTGC
GlnTyrGluCysValAsnThrAspAspSerTyrSerCysHisCys

1216 CTGAAAGGCTTTACCCTGAATCCAGATAAGAAAACCTGCAGAAGG
LeuLysGlyPheThrLeuAsnProAspLysLysThrCysArgArg

1261 ATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGC
IleAsnTyrCysAlaLeuAsnLysProGlyCysGluHisGluCys

1306 GTCAACATGGAGGAGAGCTACTACTGCCGCTGCCACCGTGGCTAC
ValAsnMetGluGluSerTyrTyrCysArgCysHisArgGlyTyr

1351 ACTCTGGACCCCAATGGCAAACCCTGCAGCCGAGTGGACCACTGT
ThrLeuAspProAsnGlyLysProCysSerArgValAspHisCys

1396 GCACAGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAG
AlaGlnGlnAspHisGlyCysGluGlnLeuCysLeuAsnThrGlu

1441 GATTCCTTCGTCTGCCAGTGCTCAGAAGGCTTCCTCATCAACGAG
AspSerPheValCysGlnCysSerGluGlyPheLeuIleAsnGlu

1486 GACCTCAAGACCTGCTCCCGGGTGGATTACTGCCTGCTGAGTGAC
AspLeuLysThrCysSerArgValAspTyrCysLeuLeuSerAsp

Fig. 5 (continued)

1531 CATGGTTGTGAATACTCCTGTGTCAACATGGACAGATCCTTTGCC
HisGlyCysGluTyrSerCysValAsnMetAspArgSerPheAla

1576 TGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAGACG
CysGlnCysProGluGlyHisValLeuArgSerAspGlyLysThr

1621 TGTGCAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAA
CysAlaLysLeuAspSerCysAlaLeuGlyAspHisGlyCysGlu

1666 CATTCGTGTGTAAGCAGTGAAGATTCGTTTGTGTGCCAGTGCTTT
HisSerCysValSerSerGluAspSerPheValCysGlnCysPhe

1711 GAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAA
GluGlyTyrIleLeuArgGluAspGlyLysThrCysArgArgLys

1756 GATGTCTGCCAAGCTATAGACCATGGCTGTGAACACATTTGTGTG
AspValCysGlnAlaIleAspHisGlyCysGluHisIleCysVal

1801 AACAGTGACGACTCATACACGTGCGAGTGCTTGGAGGGATTCCGG
AsnSerAspAspSerTyrThrCysGluCysLeuGluGlyPheArg

1846 CTCACTGAGGATGGGAAACGCTGCCGAATTTCCTCAGGGAAGGAT
LeuThrGluAspGlyLysArgCysArgIleSerSerGlyLysAsp

1891 GTCTGCAAATCAACCCACCATGGCTGCGAACACATTTGTGTTAAT
ValCysLysSerThrHisHisGlyCysGluHisIleCysValAsn

1936 AATGGGAATTCCTACATCTGCAAATGCTCAGAGGGATTTGTTCTA
AsnGlyAsnSerTyrIleCysLysCysSerGluGlyPheValLeu

1981 GCTGAGGACGGAAGACGGTGCAAGAAATGCACTGAAGGCCCAATT
AlaGluAspGlyArgArgCysLysLysCysThrGluGlyProIle

2026 GACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAG
AspLeuValPheValIleAspGlySerLysSerLeuGlyGluGlu

2071 AATTTTGAGGTCGTGAAGCAGTTTGTCACTGGAATTATAGATTCC
AsnPheGluValValLysGlnPheValThrGlyIleIleAspSer

2116 TTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGTAT
LeuThrIleSerProLysAlaAlaArgValGlyLeuLeuGlnTyr

2161 TCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCA
SerThrGlnValHisThrGluPheThrLeuArgAsnPheAsnSer

2206 GCCAAAGACATGAAAAAAGCCGTGGCCCACATGAAATACATGGGA
AlaLysAspMetLysLysAlaValAlaHisMetLysTyrMetGly

Fig. 5 (continued)

2251 AAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGA
LysGlySerMetThrGlyLeuAlaLeuLysHisMetPheGluArg

2296 AGTTTTACCCAAGGAGAAGGGGCCAGGCCCTTTTCCACAAGGGTG
SerPheThrGlnGlyGluGlyAlaArgProPheSerThrArgVal

2341 CCCAGAGCAGCCATTGTGTTCACCGACGGACGGGCTCAGGATGAC
ProArgAlaAlaIleValPheThrAspGlyArgAlaGlnAspAsp

2386 GTCTCCGAGTGGGCCAGTAAAGCCAAGGCCAATGGTATCACTATG
ValSerGluTrpAlaSerLysAlaLysAlaAsnGlyIleThrMet

2431 TATGCTGTTGGGGTAGGAAAAGCCATTGAGGAGGAACTACAAGAG
TyrAlaValGlyValGlyLysAlaIleGluGluGluLeuGlnGlu

2476 ATTGCCTCTGAGCCCACAAACAAGCATCTCTTCTATGCCGAAGAC
IleAlaSerGluProThrAsnLysHisLeuPheTyrAlaGluAsp

2521 TTCAGCACAATGGATGAGATAAGTGAAAAACTCAAGAAAGGCATC
PheSerThrMetAspGluIleSerGluLysLeuLysLysGlyIle

2566 TGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAGCA
CysGluAlaLeuGluAspSerAspGlyArgGlnAspSerProAla

2611 GGGGAACTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCA
GlyGluLeuProLysThrValGlnGlnProThrGluSerGluPro

2656 GTCACCATAAATATCCAAGACCTACTTTCCTGTTCTAATTTTGCA
ValThrIleAsnIleGlnAspLeuLeuSerCysSerAsnPheAla

2701 GTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCT
ValGlnHisArgTyrLeuPheGluGluAspAsnLeuLeuArgSer

2746 ACACAAAAGCTTTCCCATTCAACAAAACCTTCAGGAAGCCCTTTG
ThrGlnLysLeuSerHisSerThrLysProSerGlySerProLeu

2791 GAAGAAAAACACGATCAATGCAAATGTGAAAACCTTATAATGTTC
GluGluLysHisAspGlnCysLysCysGluAsnLeuIleMetPhe

2836 CAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTA
GlnAsnLeuAlaAsnGluGluValArgLysLeuThrGlnArgLeu

2881 GAAGAAATGACACAGAGAATGGAAGCCCTGGAAAATCGCCTGAGA
GluGluMetThrGlnArgMetGluAlaLeuGluAsnArgLeuArg

2926 TACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTAT
TyrArg

Fig. 5 (continued)

2971 CACGGATTACAATGAACGCAGTGCAGAGCCCCAAAGCTCAGGCTA
3016 TTGTTAAATCAATAATGTTGTGAAGTAAAACAATCAGTACTGAGA
3061 AACCTGGTTTGCCACAGAACAAAGACAAGAAGTATACACTAACTT
3106 GTATAAATTTATCTAGGAAAAAAATCCTTCAGAATTCTAAGATGA
3151 ATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAAT
3196 ATACTGTGGACACAACTTGCTTCTGCCTCATCCTGCCTTAGTGTG
3241 CAATCTCATTTGACTATACGATAAAGTTTGCACAGTCTTACTTCT
3286 GTAGAACACTGGCCATAGGAAATGCTGTTTTTTTGTACTGGACTT
3331 TACCTTGATATATGTATATGGATGTATGCATAAAATCATAGGACA
3376 TATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAAAAT
3421 TCACCACTTCAGAGAAAAGTAAAAAAA

Fig. 5 (continued)

1 CGGCCCTTCTCACACTCCTGCCCTGCTGATGTGGAACGGGGTTTG
46 GGGTTCTGCAGGGCTATTGTCTGCGCTGGGGAAGGGGACAGGCCG
91 GGACCGGGACCTCCGCTCGCAGCCGGCCGCACCAGCAGGACAGCT

136 GGCCTGAAGCTCAGAGCCGGGGCGTGCGCCATGGCCCCACACTGG
MetAlaProHisTrp

181 GCTGTCTGGCTGCTGGCAGCAAGGCTGTGGGGCCTGGGCATTGGG
AlaValTrpLeuLeuAlaAlaArgLeuTrpGlyLeuGlyIleGly

226 GCTGAGGTGTGGTGGAACCTTGTGCCGCGTAAGACAGTGTCTTCT
AlaGluValTrpTrpAsnLeuValProArgLysThrValSerSer

271 GGGGAGCTGGCCACGGTAGTACGGCGGTTCTCCCAGACCGGCATC
GlyGluLeuAlaThrValValArgArgPheSerGlnThrGlyIle

316 CAGGACTTCCTGACACTGACGCTGACGGAGCCCACTGGGCTTCTG
GlnAspPheLeuThrLeuThrLeuThrGluProThrGlyLeuLeu

361 TACGTGGGCGCCCGAGAGGCCCTGTTTGCCTTCAGCATGGAGGCC
TyrValGlyAlaArgGluAlaLeuPheAlaPheSerMetGluAla

406 CTGGAGCTGCAAGGAGCGATCTCCTGGGAGGCCCCCGTGGAGAAG
LeuGluLeuGlnGlyAlaIleSerTrpGluAlaProValGluLys

451 AAGACTGAGTGTATCCAGAAAGGGAAGAACAACCAGACCGAGTGC
LysThrGluCysIleGlnLysGlyLysAsnAsnGlnThrGluCys

496 TTCAACTTCATCCGCTTCCTGCAGCCCTACAATGCCTCCCACCTG
PheAsnPheIleArgPheLeuGlnProTyrAsnAlaSerHisLeu

541 TACGTCTGTGGCACCTACGCCTTCCAGCCCAAGTGCACCTACGTC
TyrValCysGlyThrTyrAlaPheGlnProLysCysThrTyrVal

586 AACATGCTCACCTTCACTTTGGAGCATGGAGAGTTTGAAGATGGG
AsnMetLeuThrPheThrLeuGluHisGlyGluPheGluAspGly

631 AAGGGCAAGTGTCCCTATGACCCAGCTAAGGGCCATGCTGGCCTT
LysGlyLysCysProTyrAspProAlaLysGlyHisAlaGlyLeu

676 CTTGTGGATGGTGAGCTGTACTCGGCCACACTCAACAACTTCCTG
LeuValAspGlyGluLeuTyrSerAlaThrLeuAsnAsnPheLeu

721 GGCACGGAACCCATTATCCTGCGTAACATGGGGCCCCACCACTCC
GlyThrGluProIleIleLeuArgAsnMetGlyProHisHisSer

Fig. 6

766 ATGAAGACAGAGTACCTGGCCTTTTGGCTCAACGAACCTCACTTT
MetLysThrGluTyrLeuAlaPheTrpLeuAsnGluProHisPhe

811 GTAGGCTCTGCCTATGTACCTGAGAGGGTGGGCCTGCTGTGGACA
ValGlySerAlaTyrValProGluArgValGlyLeuLeuTrpThr

856 ATGGCATACTCTCTTCCAGCCCTAGGAGGAGGGCTCCTAACAGTG
MetAlaTyrSerLeuProAlaLeuGlyGlyGlyLeuLeuThrVal

901 TAACTTATTGTGTCCCCGCGTATTTATTTGTTGTAAATATTTGAG
946 TATTTTTATATTGACAAATAAA

Fig. 6 (continued)

```
  1  GGCACCAGGCCTTCCGGAGAGACGCAGTCGGCTGCCACCCCGGGA
                                                 M

 46  TGGGTCGCTGGTGCCAGACCGTCGCGCGGGCAGCGCCCCCGGA
     etGlyArgTrpCysGlnThrValAlaArgGlyGlnArgProArgT

 91  CGTCTGCCCCCTCCCGCGCCGGTGCCCTGCTGCTGCTGCTTCTGT
     hrSerAlaProSerArgAlaGlyAlaLeuLeuLeuLeuLeuLeuL

136  TGCTGAGGTCTGCAGGTTGCTGGGGCGCAGGGGAAGCCCCGGGGG
     euLeuArgSerAlaGlyCysTrpGlyAlaGlyGluAlaProGlyA

181  CGCTGTCCACTGCTGATCCCGCCGACCAGAGCGTCCAGTGTGTCC
     laLeuSerThrAlaAspProAlaAspGlnSerValGlnCysValP

226  CCAAGGCCACCTGTCCTTCCAGCCGGCCTCGCCTTCTCTGGCAGA
     roLysAlaThrCysProSerSerArgProArgLeuLeuTrpGlnT

271  CCCCGACCACCCAGACACTGCCCTCGACCACCATGGAGACCCAAT
     hrProThrThrGlnThrLeuProSerThrThrMetGluThrGlnP

316  TCCCAGTTTCTGAAGGCAAAGTCGACCCATACCGCTCCTGTGGCT
     heProValSerGluGlyLysValAspProTyrArgSerCysGlyP

361  TTTCCTACGAGCAGGACCCCACCCTCAGGGACCCAGAAGCCGTGG
     heSerTyrGluGlnAspProThrLeuArgAspProGluAlaValA

406  CTCGGCGGTGGCCCTGGATGGTCAGCGTGCGGGCCAATGGCACAC
     laArgArgTrpProTrpMetValSerValArgAlaAsnGlyThrH

451  ACATCTGTGCCGGCACCATCATTGCCTCCCAGTGGGTGCTGACTG
     isIleCysAlaGlyThrIleIleAlaSerGlnTrpValLeuThrV

496  TGGCCCACTGCCTGATCTGGCGTGATGTTATCTACTCAGTGAGGG
     alAlaHisCysLeuIleTrpArgAspValIleTyrSerValArgV

541  TGGGGAGTCCGTGGATTGACCAGATGACGCAGACCGCCTCCGATG
     alGlySerProTrpIleAspGlnMetThrGlnThrAlaSerAspV

586  TCCCGGTGCTCCAGGTCATCATGCATAGCAGGTACCGGGCCCAGC
     alProValLeuGlnValIleMetHisSerArgTyrArgAlaGlnA

631  GGTTCTGGTCCTGGGTGGGCCAGGCCAACGACATCGGCCTCCTCA
     rgPheTrpSerTrpValGlyGlnAlaAsnAspIleGlyLeuLeuL

676  AGCTCAAGCAGGAACTCAAGTACAGCAATTACGTGCGGCCCATCT
     ysLeuLysGlnGluLeuLysTyrSerAsnTyrValArgProIleC
```

Fig. 7

```
721   GCCTGCCTGGCACGGACTATGTGTTGAAGGACCATTCCCGCTGCA
      ysLeuProGlyThrAspTyrValLeuLysAspHisSerArgCysT

766   CTGTGACGGGCTGGGGACTTTCCAAGGCTGACGGCATGTGGCCTC
      hrValThrGlyTrpGlyLeuSerLysAlaAspGlyMetTrpProG

811   AGTTCCGGACCATTCAGGAGAAGGAAGTCATCATCCTGAACAACA
      lnPheArgThrIleGlnGluLysGluValIleIleLeuAsnAsnL

856   AAGAGTGTGACAATTTCTACCACAACTTCACCAAAATCCCCACTC
      ysGluCysAspAsnPheTyrHisAsnPheThrLysIleProThrL

901   TGGTTCAGATCATCAAGTCCCAGATGATGTGTGCGGAGGACACCC
      euValGlnIleIleLysSerGlnMetMetCysAlaGluAspThrH

946   ACAGGGAGAAGTTCTGCTATGAGCTAACTGGAGAGCCCTTGGTCT
      isArgGluLysPheCysTyrGluLeuThrGlyGluProLeuValC

991   GCTCCATGGAGGGCACGTGGTACCTGGTGGGATTGGTGAGCTGGG
      ysSerMetGluGlyThrTrpTyrLeuValGlyLeuValSerTrpG

1036  GTGCAGGCTGCCAGAAGAGCGAGGCCCCACCCATCTACCTACAGG
      lyAlaGlyCysGlnLysSerGluAlaProProIleTyrLeuGlnV

1081  TCTCCTCCTACCAACACTGGATCTGGGACTGCCTCAACGGGCAGG
      alSerSerTyrGlnHisTrpIleTrpAspCysLeuAsnGlyGlnA

1126  CCCTGGCCCTGCCAGCCCCATCCAGGACCCTGCTCCTGGCACTCC
      laLeuAlaLeuProAlaProSerArgThrLeuLeuLeuAlaLeuP

1171  CACTGCCCCTCAGCCTCCTTGCTGCCCTCTGACTCTGTGTGCCCT
      roLeuProLeuSerLeuLeuAlaAlaLeu

1216  CCCTCACTTGTGGGCCCCCCTTGCCTCCGTGCCCAGGTTGCTGTG
1261  GGTGCAGCTGTCACAGCCCTGAGAGTCAGGGTGGAGATGAGGTGC
1306  TCAATTAAACATTACTGTTTTCCATGTAAAAAAAAAAAAAAAAAA
1351  AAAAAAAAA
```

Fig. 7 (continued)

CACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCCCCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACC
81
AGTGATGGTAAGTGCTGGCCCAGACTGAAGCTCGGAGAGGCACTCTGCTTGCCCAGCGTCACAGTCTTAGCTCCCAACTG
161
TCCTGGCTTCCAGTCTCCCTTGCTTCCCAGATCCCAGACTCTAGCCCCAGCCCCGTCTCTTTCACCAGCTCCTGGGACCC
241
TACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAACTGTATCTACAACCAGCTGCACCAGCGACACCTGTCC
321
AACCCGGCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTGGGGTGCAGGGCCCCTGTCAGGTCTGATAGGGAGAAGA
401
GAAGGAGCAGAAGGGGAGGGGCCTAACCCTGGGCTGGGGGTTGGACTCACAGGACTGGGGGAAAGAGCTGCAATCAGAGG
481
GTGTCTGCCATAGCTGGGCTCAGGCATCTGTCCTTGGCTTTGTTGCCTGGCTCCAGGGAGATTCCGGGGGCCCTGTGCTG
561
TGCCTCGAGCCTGACGGACACTGGGTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGT
641
GCTGCTGACCAACACAGCTGCTCACAGTTCCTGGCTGCAGGCTCGAGTTCAGGGGGCAGCTTTCCTGGCCCAGAGCCCAG
721
AGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGATCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCC
MetSerAspGluAspSerCysValAlaCysGlySerLeuArgThrAlaGlyProGlnAlaGlyAlaPro
801
TCCCCATGGCCCTGGGAGGCCAGGCTGATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTCAGAGGAGGCGGT
SerProTrpProTrpGluAlaArgLeuMetHisGlnGlyGlnLeuAlaCysGlyGlyAlaLeuValSerGluGluAlaVa
881
GCTAACTGCTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGACCGGAGGAGT
lLeuThrAlaAlaHisCysPheIleGlyArgGlnAlaProGluGluTrpSerValGlyLeuGlyThrArgProGluGluT
961
GGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGGCTACGACATGGCCCTCCTGCTGCTGGCC
rpGlyLeuLysGlnLeuIleLeuHisGlyAlaTyrThrHisProGluGlyGlyTyrAspMetAlaLeuLeuLeuLeuAla
1041
CAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCCTCTGCCTGCCCTATGCTGACCACCACCTGCCTGATGGGGAGCGTGG
GlnProValThrLeuGlyAlaSerLeuArgProLeuCysLeuProTyrAlaAspHisHisLeuProAspGlyGluArgGl
1121
CTGGGTTCTGGGACGGGCCCGCCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAGGG
yTrpValLeuGlyArgAlaArgProGlyAlaGlyIleSerSerLeuGlnThrValProValThrLeuLeuGlyProArgA
1201
CCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGGTGTGTACCAGTGCTGTGGGT
laCysSerArgLeuHisAlaAlaProGlyGlyAspGlySerProIleLeuProGlyMetValCysThrSerAlaValGly
1281
GAGCTGCCCAGCTGTGAGGTGAGCCCCAGGCCCCCACACCTTACCTAACAGGCCCCTGGCATCCCCTCACCCAATAGCTC
GluLeuProSerCysGluValSerProArgProProHisLeuThr
1361
AAGAACGGACCTTCCAGGCTTGGCCTCTGGACCCACCTCCCACCTGAAGCTAAGCCTTTTTGCCAATTAGCCCCCAAACA
1441
GCCAG

Fig. 8

```
1    CTTAACAGCCACTTGTTTCATCCCACCTGGGCATTAGGTTGACTT

46   CAAAGATGCCTCAGTTACTGCAAAACATTAATGGGATCATCGAGG
          MetProGlnLeuLeuGlnAsnIleAsnGlyIleIleGluA

91   CCTTCAGGCGCTATGCAAGGACGGAGGGCAACTGCACAGCGCTCA
     laPheArgArgTyrAlaArgThrGluGlyAsnCysThrAlaLeuT

136  CCCGAGGGGAGCTGAAAAGACTCTTGGAGCAAGAGTTTGCCGATG
     hrArgGlyGluLeuLysArgLeuLeuGluGlnGluPheAlaAspV

181  TGATTGTGAAACCCCACGATCCAGCAACTGTGGATGAGGTCCTGC
     alIleValLysProHisAspProAlaThrValAspGluValLeuA

226  GTCTGCTGGATGAAGACCACACAGGGACTGTGGAATTCAAGGAAT
     rgLeuLeuAspGluAspHisThrGlyThrValGluPheLysGluP

271  TCCTGGTCTTAGTGTTTAAAGTTGCCCAGGCCTGTTTCAAGACAC
     heLeuValLeuValPheLysValAlaGlnAlaCysPheLysThrL

316  TGAGCGAGAGTGCTGAGGGAGCCTGCGGCTCTCAAGAGTCTGGAA
     euSerGluSerAlaGluGlyAlaCysGlySerGlnGluSerGlyS

361  GCCTCCACTCTGGGGCCTCGCAGGAGCTGGGCGAAGGACAGAGAA
     erLeuHisSerGlyAlaSerGlnGluLeuGlyGluGlyGlnArgS

406  GTGGCACTGAAGTGGGAAGGGCGGGGAAAGGGCAGCATTATGAGG
     erGlyThrGluValGlyArgAlaGlyLysGlyGlnHisTyrGluG

451  GGAGCAGCCACAGACAGAGCCAGCAGGGTTCCAGAGGGCAGAACA
     lySerSerHisArgGlnSerGlnGlnGlySerArgGlyGlnAsnA

496  GGCCTGGGGTTCAGACCCAGGGTCAGGCCACTGGCTCTGCGTGGG
     rgProGlyValGlnThrGlnGlyGlnAlaThrGlySerAlaTrpV

541  TCAGCAGCTATGACAGGCAAGCTGAGTCCCAGAGCCAGGAAAGAA
     alSerSerTyrAspArgGlnAlaGluSerGlnSerGlnGluArgI

586  TAAGCCCGCAGATACAACTCTCTGGGCAGACAGAGCAGACCCAGA
     leSerProGlnIleGlnLeuSerGlyGlnThrGluGlnThrGlnL

631  AAGCTGGAGAAGGCAAGAGGAATCAGACAACAGAGATGAGGCCAG
     ysAlaGlyGluGlyLysArgAsnGlnThrThrGluMetArgProG

676  AGAGACAGCCACAGACCAGGGAACAGGACAGAGCCCACCAGACAG
     luArgGlnProGlnThrArgGluGlnAspArgAlaHisGlnThrG
```

Fig. 9

[redacted]

```
721   GTGAGACTGTGACTGGATCTGGAACTCAGACCCAGGCAGGTGCCA
      lyGluThrValThrGlySerGlyThrGlnThrGlnAlaGlyAlaT

766   CCCAGACTGTGGAGCAGGACAGCAGCCACCAGACAGGAAGCACCA
      hrGlnThrValGluGlnAspSerSerHisGlnThrGlySerThrS

811   GCACCCAGACACAGGAGTCCACCAATGGCCAGAACAGAGGGACTG
      erThrGlnThrGlnGluSerThrAsnGlyGlnAsnArgGlyThrG

856   AGATCCACGGTCAAGGCAGGAGCCAGACCAGCCAGGCTGTGACAG
      luIleHisGlyGlnGlyArgSerGlnThrSerGlnAlaValThrG

901   GAGGACACACTCAGATACAGGCAGGGTCACACACCGAGACTGTGG
      lyGlyHisThrGlnIleGlnAlaGlySerHisThrGluThrValG

946   AGCAGGACAGAAGCCAAACTGTAAGCCACGGAGGGGCTAGAGAAC
      luGlnAspArgSerGlnThrValSerHisGlyGlyAlaArgGluG

991   AGGGACAGACCCAGACGCAGCCAGGCAGTGGTCAAAGATGGATGC
      lnGlyGlnThrGlnThrGlnProGlySerGlyGlnArgTrpMetG

1036  AAGTGAGCAACCCTGAGGCAGGAGAGACAGTACCGGGAGGACAGG
      lnValSerAsnProGluAlaGlyGluThrValProGlyGlyGlnA

1081  CCCAGACTGGGGCAAGCACTGAGTCAGGAAGGCAGGAGTGGAGCA
      laGlnThrGlyAlaSerThrGluSerGlyArgGlnGluTrpSerS

1126  GCACTCACCCAAGGCGCTGTGTGACAGAAGGGCAGGGAGACAGAC
      erThrHisProArgArgCysValThrGluGlyGlnGlyAspArgG

1171  AGCCCACAGTGGTTGGTGAGGAATGGGTTGATGACCACTCAAGGG
      lnProThrValValGlyGluGluTrpValAspAspHisSerArgG

1216  AGACAGTGATCCTCAGGCTGGACCAGGGCAACTTGCATACCAGTG
      luThrValIleLeuArgLeuAspGlnGlyAsnLeuHisThrSerV

1261  TTTCCTCAGCACAGGGCCAGGATGCAGCCCAGTCAGAAGAGAAGC
      alSerSerAlaGlnGlyGlnAspAlaAlaGlnSerGluGluLysA

1306  GAGGCATCACAGCTAGAGAGCTGTATTCCTACTTGAGAAGCACCA
      rgGlyIleThrAlaArgGluLeuTyrSerTyrLeuArgSerThrL

1351  AGCCATGACTTCCCCGACTCCAATGTCCAGTACTGGAAGAAGACA
      ysPro

1396  GCTGGAGAGAGTTTGGCTTGTCCTGCATGGCCAATCCAGTGGGTG
1441  CATCCCTGGACATCAGCTCTTCATTATGCAGCTTCCCTTTTAGGT
1486  CTTTCTCAATGAGATAATTTCTGCAAGGAGCTTTCTATCCTGAAC
1531  TCTTCTTTCTTACCTGCTTTGCGGTGCAGACCCTCTCAGGAGCAG
1576  GAAGACTCAGAACAAGTCACCCCTT
```

Fig. 9 (continued)

1. 11618130.0.184_Cura_108
2. 11618130.0.27_Cura_56

```
116181300184_cura_108  MSDEDSCVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGR
11618130027_cura_56    MSDEDSCVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFNGR

116181300184_cura_108  QAPEEWSVGLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYA
11618130027_cura_56    QAPEEWSVGLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRALCLPYF

116181300184_cura_108  DHHLPDGERGWVLGRARPGAGISSLQTVPVTLLGPRACSRLHAAPGGDGSPILPGMVCTS
11618130027_cura_56    DHHLPDGERGWVLGRARPGAGISSLQTVPVTLLGPRACSRLHAAPGGDGSPILPGMVCTS

116181300184_cura_108  AVGELPSCEVSPRPPHLT~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
11618130027_cura_56    AVGELPSCEGLSGAPLVHEVRGTWFLAGLHSFGDACQGPARPAVFTALPAYEDWVSSLDW

116181300184_cura_108  ~~~~~~~~~~~~~~~~~~~~~~~~~~~
11618130027_cura_56    QVYFAEEPEPEAEPGSCLANISQPTSC
```

Fig. 10

10074566.021302

Sequences analyzed:

1. **14578444-0-47_Cura_56**
2. **14578444-0-143Cura_56**

```
14578444047_cura_56     MEKMLAGCFLLILGQIVLLPCEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVF
145784440143_cura_56    MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVF

14578444047_cura_56     IIDSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEV
145784440143_cura_56    IIDSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEV

14578444047_cura_56     ERAVKRMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEV
145784440143_cura_56    ERAVKRMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEV

14578444047_cura_56     AAKARDTGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAH
145784440143_cura_56    AAKARDTGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAH

14578444047_cura_56     MCSTLEHNCAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGS
145784440143_cura_56    MCSTLEHNCAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGS

14578444047_cura_56     FVCECYSGYALAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCT
145784440143_cura_56    FVCECYSGYALAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCT
```

Fig. 11

```
14578444047_cura_56   KIDYCASSNHGCQYECVNTDDSYSCHCLKGFTLNPDKKTCRRINYCALNKPGCEHECVNM
145784440143_cura_56  KIDYCASSNHGCQYECVNTDDSYSCHCLKGFTLNPDKKTCRRINYCALNKPGCEHECVNM

14578444047_cura_56   EESYYCRCHRGYTLDPNGKPCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLINEDLK
145784440143_cura_56  EESYYCRCHRGYTLDPNGKPCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLINEDLK

14578444047_cura_56   TCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCEHSC
145784440143_cura_56  TCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCEHSC

14578444047_cura_56   VSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLTE
145784440143_cura_56  VSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLTE

14578444047_cura_56   DGKRCRISSGKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLV
145784440143_cura_56  DGKRCRISSGKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLV

14578444047_cura_56   FVIDGSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKD
145784440143_cura_56  FVIDGSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKD

14578444047_cura_56   MKKAVAHMKYMGKGSMTGLALKHMFERSFTQGEGARPFSTRVPRAAIVFTDGRAQDDVSE
145784440143_cura_56  MKKAVAHMKYMGKGSMTGLALKHMFERSFTQGEGARPLFHKGAQSSHCVHRRTGSG~~~~

14578444047_cura_56   WASKAKANGITMYAVGVGKAIEEELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEA
145784440143_cura_56  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

14578444047_cura_56   LEDSDGRQDSPAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQK
145784440143_cura_56  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

14578444047_cura_56   LSHSTKPSGSPLEEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR
145784440143_cura_56  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

Fig. 11 (continued)

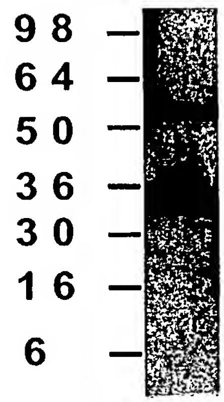

Fig. 12

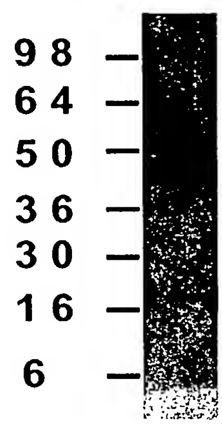

Fig. 13

Normal & Tumor Tissues	11696905	16406477.0.206	21433858	21637262.0.64
Endothelial cells	3.5	0.0	6.6	0.0
Endothelial cells (treated)	2.9	0.0	2.0	0.0
Pancreas	9.4	3.1	1.2	0.0
Pancreatic ca. CAPAN 2	3.7	0.0	0.3	0.0
Adipose	60.7	0.3	22.5	0.8
Adrenal gland	18.0	0.0	3.2	0.2
Thyroid	13.8	0.0	4.6	2.4
Salivary gland	0.0	0.6	0.7	36.3
Pituitary gland	2.2	0.6	4.0	1.4
Brain (fetal)	3.1	0.5	6.9	0.7
Brain (whole)	4.4	0.7	24.5	0.3
Brain (amygdala)	17.2	0.1	5.0	0.4
Brain (cerebellum)	1.6	1.2	41.8	1.4
Brain (hippocampus)	9.3	0.8	10.4	0.6
Brain (hypothalamus)	5.7	10.0	2.3	0.5
Brain (substantia nigra)	33.2	0.7	5.2	0.1
Brain (thalamus)	22.7	0.5	5.2	0.0
Spinal cord	21.8	0.3	4.0	1.5
CNS ca. (glio/astro) U87-MG	2.2	0.0	1.0	0.0
CNS ca. (glio/astro) U-118-MG	4.5	0.0	1.5	0.0
CNS ca. (astro) SW1783	0.0	0.0	0.7	0.0
CNS ca.* (neuro; met) SK-N-AS	2.7	0.0	12.6	0.1
CNS ca. (astro) SF-539	0.2	0.0	0.0	0.0
CNS ca. (astro) SNB-75	1.3	0.0	0.6	0.0
CNS ca. (glio) SNB-19	0.6	0.0	0.8	0.2
CNS ca. (glio) U251	0.2	0.0	3.6	0.1
CNS ca. (glio) SF-295	6.2	0.1	0.2	0.0
Heart	10.7	0.1	1.3	0.1
Skeletal muscle	18.4	0.0	0.2	0.2
Bone marrow	11.1	0.0	0.1	0.0
Thymus	7.3	0.9	2.5	0.5
Spleen	2.9	0.1	1.4	0.0
Lymph node	4.3	0.1	1.3	0.1

Fig. 14

Normal & Tumor Tissues	11696905	16406477.0.206	21433858	21637262.0.64
Colon (ascending)	1.3	0.2	5.1	1.3
Stomach	5.4	0.2	5.7	0.0
Small intestine	7.0	0.2	1.7	0.0
Colon ca. SW480	0.4	0.0	0.0	0.1
Colon ca.* (SW480 met)SW620	0.1	0.0	0.0	0.0
Colon ca. HT29	0.4	0.0	0.0	0.1
Colon ca. HCT-116	4.4	0.0	0.0	0.0
Colon ca. CaCo-2	1.1	0.1	0.1	0.0
Colon ca. HCT-15	11.0	0.2	0.3	0.2
Colon ca. HCC-2998	0.0	0.0	1.3	0.0
Gastric ca.* (liver met) NCI-N87	4.9	0.3	1.9	0.0
Bladder	18.8	0.1	10.8	0.1
Trachea	4.8	0.0	2.2	100.0
Kidney	7.3	0.4	13.1	0.1
Kidney (fetal)	11.0	1.8	29.5	0.1
Renal ca. 786-0	0.4	0.0	0.0	0.0
Renal ca. A498	56.3	0.0	0.0	0.1
Renal ca. RXF 393	2.7	0.0	0.1	0.0
Renal ca. ACHN	1.0	0.0	0.1	0.1
Renal ca. UO-31	1.8	0.0	0.4	0.1
Renal ca. TK-10	13.4	0.5	0.2	0.1
Liver	74.7	0.7	2.1	0.1
Liver (fetal)	27.7	1.2	0.9	0.0
Liver ca. (hepatoblast HepG2	7.4	0.0	0.0	0.0
Lung	9.9	0.0	2.9	0.0
Lung (fetal)	1.5	1.5	3.0	0.0
Lung ca. (small cell) LX-1	0.4	0.0	0.0	0.0
Lung ca. (small cell) NCI-H69	0.5	0.1	9.3	0.5
Lung ca. (s.cell var.) SHP-77	0.6	0.4	100.0	1.7
Lung ca. (large cell) NCI-H460	20.6	0.3	66.9	0.6
Lung ca. (non-sm. cell) A549	3.3	0.0	15.5	0.1
Lung ca. (non-s.cell) NCI-H23	7.4	0.5	9.0	0.0
Lung ca (non-s.cell) HOP-62	32.1	0.1	1.5	0.1
Lung ca. (non-s.cl) NCI-H522	11.0	0.6	0.0	0.0
Lung ca. (squam.) SW 900	3.3	0.9	6.1	0.1

Fig. 14 (continued)

Normal & Tumor Tissues	11696905.0	16406477.0.206	21433858.0	21637262.0.64
Mammary gland	30.4	1.5	12.2	0.0
Breast ca.* (pl. effusion) MCF-7	4.8	0.0	0.0	0.0
Breast ca.* (pl.ef) MDA-MB-231	2.2	0.0	0.0	0.1
Breast ca.* (pl. effusion) T47D	9.8	0.1	0.9	0.6
Breast ca. BT-549	9.2	0.1	1.2	0.3
Breast ca. MDA-N	1.3	0.0	0.0	0.0
Ovary	6.0	0.3	9.7	0.0
Ovarian ca. OVCAR-3	1.6	0.1	0.8	0.1
Ovarian ca. OVCAR-4	1.9	0.0	0.0	0.0
Ovarian ca. OVCAR-5	7.1	0.3	6.9	0.6
Ovarian ca. OVCAR-8	1.3	2.7	2.7	0.0
Ovarian ca. IGROV-1	0.7	0.2	5.0	0.0
Ovarian ca.* (ascites) SK-OV-3	2.5	0.0	0.2	0.0
Myometrium	2.3	0.0	41.2	1.2
Uterus	6.3	0.6	25.7	0.1
Placenta	100.0	0.0	94.0	0.1
Prostate	13.3	0.1	3.4	0.1
Prostate ca.* (bone met) PC-3	7.9	1.7	0.2	0.2
Testis	14.3	100.0	37.1	4.0
Melanoma Hs688(A).T	1.4	0.0	0.0	0.0
Melanoma* (met) Hs688(B).T	5.3	0.0	0.0	0.0
Melanoma UACC-62	0.6	0.0	0.0	0.0
Melanoma M14	0.9	0.1	0.3	0.2
Melanoma LOX IMVI	1.0	0.0	0.0	0.1
Melanoma* (met) SK-MEL-5	0.0	0.0	8.7	0.0
Melanoma SK-MEL-28	100.0	0.0	0.0	0.0

Fig. 14 (continued)

Drawings

Figure 15. Nucleotide Sequence for CG106318-01.

>CG106318-01 4810 nt
GTCCATGGGGCCGATGTATGGGAGATGAATGTGGTCCCGGAGGCATCCAAACGAGGGCTG
TGTGGTGTGCTCATGTGGAGGGATGGACTACACTGCATACTAACTGTAAGCAGGCCGAGA
GACCCAATAACCAGCAGAATTGTTTCAAAGTTTGCGATTGGCACAAAGAGTTGTACGACT
GGAGACTGGGACCTTGGAATCAGTGTCAGCCCGTGATTTCAAAAAGCCTAGAGAAACCTC
TTGAGTGCATTAAGGGGGAAGAAGGTATTCAGGTGAGGGAGATAGCGTGCATCCAGAAAG
ACAAAGACATTCCTGCGGAGGATATCATCTGTGAGTACTTTGAGCCCAAGCCTCTCCTGG
AGCAGGCTTGCCTCATTCCTTGCCAGCAAGATTGCATCGTGTCTGAATTTTCTGCCTGGT
CCGAATGCTCCAAGACCTGCGGCAGCGGGCTCCAGCACCGGACGCGTCATGTGGTGGCGC
CCCCGCAGTTCGGAGGCTCTGGCTGTCCAAACCTGACGGAGTTCCAGGTGTGCCAATCCA
GTCCATGCGAGGCCGAGGAGCTCAGGTACAGCCTGCATGTGGGGCCCTGGAGCACCTGCT
CAATGCCCCACTCCCGACAAGTAAGACAAGCAAGGAGACGCGGGAAGAATAAAGAACGGG
AAAAGGACCGCAGCAAAGGAGTAAAGGATCCAGAAGCCCGCGAGCTTATTAAGAAAAAGA
GAAACAGAAACAGGCAGAACAGACAAGAGAACAAATATTGGGACATCCAGATTGGATATC
AGACCAGAGAGGTTATGTGCATTAACAAGACGGGGAAAGCTGCTGATTTAAGCTTTTGCC
AGCAAGAGAAGCTTCCAATGACCTTCCAGTCCTGTGTGATCACCAAAGAGTGCCAGGTTT
CCGAGTGGTCAGAGTGGAGCCCCTGCTCAAAAACATGCCATGACATGGTGTCCCCTGCAG
GCACTCGTGTAAGGACACGAACCATCAGGCAGTTTCCCATTGGCAGTGAAAAGGAGTGTC
CAGAATTTGAAGAAAAAGAACCCTGTTTGTCTCAAGGAGATGGAGTTGTCCCCTGTGCCA
CGTATGGCTGGAGAACTACAGAGTGGACTGAGTGCCGTGTGGACCCTTTGCTCAGTCAGC
AGGACAAGAGGCGCGGCAACCAGACGGCCCTCTGTGGAGGGGGCATCCAGACCCGAGAGG
TGTACTGCGTGCAGGCCAACGAAAACCTCCTCTCACAATTAAGTACCCACAAGAACAAAG
AAGCCTCAAAGCCAATGGACTTAAAATTATGCACTGGACCTATCCCTAATACTACACAGC
TGTGCCACATTCCTTGTCCAACTGAATGTGAAGTTTCACCTTGGTCAGCTTGGGGACCTT
GTACTTATGAAAACTGTAATGATCAGCAAGGGAAAAAAGGCTTCAAACTGAGGAAGCGGC
GCATTACCAATGAGCCCACTGGAGGCTCTGGGGTAACCGGAAACTGCCCTCACTTACTGG
AAGCCATTCCCTGTGAAGAGCCTGCCTGTTATGACTGGAAAGCGGTGAGACTGGGAGACT
GCGAGCCAGATAACGGAAAGGAGTGTGGTCCAGGCACGCAAGTTCAAGAGGTTGTGTGCA
TCAACAGTGATGGAGAAGAAGTTGACAGACAGCTGTGCAGAGATGCCATCTTCCCCATCC
CTGTGGCCTGTGATGCCCCATGCCCGAAAGACTGTGTGCTCAGCACATGGTCTACGTGGT
CCTCCTGCTCACACACCTGCTCAGGGAAAACGACAGAAGGGAAACAGATACGAGCACGAT
CCATTCTGGCCTATGCGGGTGAAGAAGGTGGAATTCGCTGTCCAAATAGCAGTGCTTTGC
AAGAAGTACGAAGCTGTAATGAGCATCCTTGCACAGTGTACCACTGGCAAACTGGTCCCT
GGGGCCAGTGCATTGAGGACACCTCAGTATCGTCCTTCAACACAACTACGACTTGGAATG
GGGAGGCCTCCTGCTCTGTCGGCATGCAGACAAGAAAAGTCATCTGTGTGCGAGTCAATG
TGGGCCAAGTGGGACCCAAAAAATGTCCTGAAAGCCTTCGACCTGAAACTGTAAGGCCTT
GTCTGCTTCCTTGTAAGAAGGACTGTATTGTGACCCCATATAGTGACTGGACATCATGCC
CCTCTTCGTGTAAAGAAGGGGACTCCAGTATCAGGAAGCAGTCTAGGCATCGGGTCATCA
TTCAGCTGCCAGCCAACGGGGGCCGAGACTGCACAGATCCCCTCTATGAAGAGAAGGCCT
GTGAGGCACCTCAAGCGTGCCAAAGCTACAGGTGGAAGACTCACAAATGGCGCAGATGCC
AATTAGTCCCTTGGAGCGTGCAACAAGACAGCCCTGGAGCACAGGAAGGCTGTGGGCCTG
GGCGACAGGCAAGAGCCATTACTTGTCGCAAGCAAGATGGAGGACAGGCTGGAATCCATG
AGTGCCTACAGTATGCAGGCCCTGTGCCAGCCCTTACCCAGGCCTGCCAGATCCCCTGCC
AGGATGACTGTCAATTGACCAGCTGGTCCAAGTTTTCTTCATGCAATGGAGACTGTGGTG
CAGTTAGGACCAGAAAGCGCACTCTTGTTGGAAAAAGTAAAAAGAAGGAAAAATGTAAAA
ATTCCCATTTGTATCCCCTGATTGAGACTCAGTATTGTCCTTGTGACAAATATAATGCAC
AACCTGTGGGGAACTGGTCAGACTGTATTTTACCAGAGGGAAAAGTGGAAGTGTTGCTGG
GAATGAAAGTACAAGGAGACATCAAGGAATGCGGACAAGGATATCGTTACCAAGCAATGG
CATGCTACGATCAAAATGGCAGGCTTGTGGAAACATCTAGATGTAACAGCCATGGTTACA
TTGAGGAGGCCTGCATCATCCCCTGCCCCTCAGACTGCAAGCTCAGTGAGTGGTCCAACT
GGTCGCGCTGCAGCAAGTCCTGTGGGAGTGGTGTGAAGGTTCGTTCTAAATGGCTGCGTG
AAAAACCATATAATGGAGGAAGGCCTTGCCCCAAACTGGACCATGTCAACCAGGCACAGG
TGTATGAGGTTGTCCCATGCCACAGTGACTGCAACCAGTACCTATGGGTCACAGAGCCCT
GGAGCATCTGCAAGGTGACCTTTGTGAATATGCGGGAGAACTGTGGAGAGGGCGTGCAAA
CCCGAAAAGTGAGATGCATGCAGAATACAGCAGATGGCCCTTCTGAACATGTAGAGGATT
ACCTCTGTGACCCAGAAGAGATGCCCCTGGGCTCTAGAGTGTGCAAATTACCATGCCCTG
AGGACTGTGTGATATCTGAATGGGGTCCATGGACCCAATGTGTTTTGCCTTGCAATCAAA
GCAGTTTCCGGCAAAGGTCAGCTGATCCCATCAGACAACCAGCTGATGAAGGAAGATCTT
GCCCTAATGCTGTTGAGAAAGAACCCTGTAACCTGAACAAAAACTGCTACCACTATGATT
ATAATGTAACAGACTGGAGTACATGTCAGCTGAGTGAGAAGGCAGTTTGTGGAAATGGAA
TAAAAACAAGGATGTTGGATTGTGTTCGAAGTGATGGCAAGTCAGTTGACCTGAAATATT

GTGAAGCGCTTGGCTTGGAGAAGAACTGGCAGATGAACACGTCCTGCATGGTGGAATGCC
CTGTGAACTGTCAGCTTTCTGATTGGTCTCCTTGGTCAGAATGTTCTCAAACATGTGGCC
TCACAGGAAAAATGATCCGAAGACGAACAGTGACCCAGCCCTTTCAAGGTGATGGAAGAC
CATGCCCTTCCCTGATGGACCAGTCCAAACCCTGCCCAGTGAAGCCTTGTTATCGGTGGC
AATATGGCCAGTGGTCTCCATGCCAAGTGCAGGAGGCCCAGTGTGGAGAAGGGACCAGAA
CAAGGAACATTTCTTGTGTAGTAAGTGATGGGTCAGCTGATGATTTCAGCAAAGTGGTGG
ATGAGGAATTCTGTGCTGACATTGAACTCATTATAGATGGTAATAAAAATATGGTTCTGG
AGGAATCCTGCAGCCAGCCTTGCCCAGGTGACTGTTATTTGAAGGACTGGTCTTCCTGGA
GCCTGTGTCAGCTGACCTGTGTGAATGGTGAGGATCTAGGCTTTGGTGGAATACAGGTCA
GATCCAGACCGGTGATTATACAAGAACTAGAGAATCAGCATCTGTGCCCAGAGCAGATGT
TAGAAACAAAATCATGTTATGATGGACAGTGCTATGAATATAAATGGATGGCCAGTGCTT
GGAAGGGCTCTTCCCGAACAGTGTGGTGTCAAAGGTCAGATGGTATAAATGTAACAGGGG
GCTGCTTGGTGATGAGCCAGCCTGATGCCGACAGGTCTTGTAACCCACCGTGTAGTCAAC
CCCACTCGTACTGTAGCGAGACAAAAACATGCCATTGTGAAGAAGGGTACACTGAAGTCA
TGTCTTCTAACAGCACCCTTGAGCAATGCACACTTATCCCCGTGGTGGTATTACCCACCA
TGGAGGACAAAAGAGGAGATGTGAAAACCAGTCGGGCTGTACATCCAACCCAACCCTCCA
GTAACCCAGCAGGACGGGGAAGGACCTGGTTTCTACAGCCATTTGGGCCAGATGGGAGAC
TAAAGACCTGGGTTTACGGTGTAGCAGCTGGGGCATTTGTGTTACTCATCTTTATTGTCT
CCATGATTTATCTAGCTTGCAAAAAGCCAAAGAAACCCCAAAGAAGGCAAAACAACCGAC
TGAAACCTTTAACCTTAGCCTATGATGGAGATGCCGACATGTAACATATAACTTTTCCTG
GCAACAACCA (SEQ ID NO: 40)

Protein Sequence for CG106318-01 ORF Start: 18 ORF Stop: 4782 Frame: 3

Protein Sequence:

>CG106318-01-prot 1588 aa
MGDECGPGGIQTRAVWCAHVEGWTTLHTNCKQAERPNNQQNCFKVCDWHKELYDWRLGPW
NQCQPVISKSLEKPLECIKGEEGIQVREIACIQKDKDIPAEDIICEYFEPKPLLEQACLI
PCQQDCIVSEFSAWSECSKTCGSGLQHRTRHVVAPPQFGGSGCPNLTEFQVCQSSPCEAE
ELRYSLHVGPWSTCSMPHSRQVRQARRRGKNKEREKDRSKGVKDPEARELIKKKRNRNRQ
NRQENKYWDIQIGYQTREVMCINKTGKAADLSFCQQEKLPMTFQSCVITKECQVSEWSEW
SPCSKTCHDMVSPAGTRVRTRTIRQFPIGSEKECPEFEEKEPCLSQGDGVVPCATYGWRT
TEWTECRVDPLLSQQDKRRGNQTALCGGGIQTREVYCVQANENLLSQLSTHKNKEASKPM
DLKLCTGPIPNTTQLCHIPCPTECEVSPWSAWGPCTYENCNDQQGKKGFKLRKRRITNEP
TGGSGVTGNCPHLLEAIPCEEPACYDWKAVRLGDCEPDNGKECGPGTQVQEVVCINSDGE
EVDRQLCRDAIFPIPVACDAPCPKDCVLSTWSTWSSCSHTCSGKTTEGKQIRARSILAYA
GEEGGIRCPNSSALQEVRSCNEHPCTVYHWQTGPWGQCIEDTSVSSFNTTTTWNGEASCS
VGMQTRKVICVRVNVGQVGPKKCPESLRPETVRPCLLPCKKDCIVTPYSDWTSCPSSCKE
GDSSIRKQSRHRVIIQLPANGGRDCTDPLYEEKACEAPQACQSYRWKTHKWRRCQLVPWS
VQQDSPGAQEGCGPGRQARAITCRKQDGGQAGIHECLQYAGPVPALTQACQIPCQDDCQL
TSWSKFSSCNGDCGAVRTRKRTLVGKSKKKEKCKNSHLYPLIETQYCPCDKYNAQPVGNW
SDCILPEGKVEVLLGMKVQGDIKECGQGYRYQAMACYDQNGRLVETSRCNSHGYIEEACI
IPCPSDCKLSEWSNWSRCSKSCGSGVKVRSKWLREKPYNGGRPCPKLDHVNQAQVYEVVP
CHSDCNQYLWVTEPWSICKVTFVNMRENCGEGVQTRKVRCMQNTADGPSEHVEDYLCDPE
EMPLGSRVCKLPCPEDCVISEWGPWTQCVLPCNQSSFRQRSADPIRQPADEGRSCPNAVE
KEPCNLNKNCYHYDYNVTDWSTCQLSEKAVCGNGIKTRMLDCVRSDGKSVDLKYCEALGL
EKNWQMNTSCMVECPVNCQLSDWSPWSECSQTCGLTGKMIRRRTVTQPFQGDGRPCPSLM
DQSKPCPVKPCYRWQYGQWSPCQVQEAQCGEGTRTRNISCVVSDGSADDFSKVVDEEFCA
DIELIIDGNKNMVLEESCSQPCPGDCYLKDWSSWSLCQLTCVNGEDLGFGGIQVRSRPVI
IQELENQHLCPEQMLETKSCYDGQCYEYKWMASAWKGSSRTVWCQRSDGINVTGGCLVMS
QPDADRSCNPPCSQPHSYCSETKTCHCEEGYTEVMSSNSTLEQCTLIPVVVLPTMEDKRG
DVKTSRAVHPTQPSSNPAGRGRTWFLQPFGPDGRLKTWVYGVAAGAFVLLIFIVSMIYLA
CKKPKKPQRRQNNRLKPLTLAYDGDADM (SEQ ID NO: 41)

Figure 16. Nucleotide and Protein Sequences for CG50817-04.

>CG50817-04 1447 nt
GCGGACACCAGTGATGCTCCTGGGACCCTACGCAATCTGCGCCTGCGTCTCATCAGTCGC
CCCACATGTAACTGTATCTACAACCAGCTGCACCAGCGACACCTGTCCAACCCGGCCCGG
CCTGGGATGCTATGTGGGGGCCCCCAGCCTGGGGTGCAGGGCCCCTGTCAGGTCTGATAG
GGAGAAGAGAAGGAGCAGAAGGGGAGGGGCCTAACCCTGGGCTGGGGGTTGGACTCACAG
GACTGGGGGAAAGAGCTGCAATCAGAGGGTGTCTGCCATAGCTGGGCTCAGGCATCTGTC
CTTGGCTTTGTTGCCTGGCTCCAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCT
GACGGACACTGGGTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGAC
GCTCCTGTGCTGCTGACCAACACAGCTGCTCACAGTTCCTGGCTGCAGGCTCGAGTTCAG
GGGGCAGCTTTCCTGGCCCAGAGCCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGT
GTAGCCTGTGGATCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCCTCCCCATGGCCC
TGGGAGGCCAGGCTGATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTCAGAG
GAGGCGGTGCTAACTGCTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGC
GTAGGGCTGGGGACCAGACCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCC
TACACCCACCCTGAGGGGGGCTACGACATGGCCCTCCTGCTGCTGGCCCAGCCTGTGACA
CTGGGAGCCAGCCTGCGGCCCCTCTGCCTGCCCTATGCTGACCACCACCTGCCTGATGGG
GAGCGTGGCTGGGTTCTGGGACGGGCCCGCCCAGGAGCAGGCATCAGCTCCCTCCAGACA
GTGCCCGTGACCCTCCTGGGGCCTAGGGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGT
GATGGCAGCCCTATTCTGCCGGGGATGGTGTGTACCAGTGCTGTGGGTGAGCTGCCCAGC
TGTGAGGCCAACCAACCAGCTGCTGACAGGGGACCTGGCCATTCTCAGGAACAAGAGAAT
GCAGGCAGGCAAATGGCATTACTGCCCCTGTCCTCCCCACCCTGTCATGTGTGATTCCAG
GCACCAGGGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGG
TGCCCACTCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCCACACCCAACTCTGC
TACCAAGCAGGCGTCTCAGCTTTCCTCCTCCTTTACCCTTTCAGATACAATCACGCCAGC
CACGTTGTTTTGAAAATTTCTTTTTTTGGGGGGCAGCAGTTTTCCTTTTTTTAAACTTAA
ATAAATT (SEQ ID NO:42)

Protein Sequence for CG50817-04 ORF Start: 520 ORF Stop: 1192 Frame: 1

Protein Sequence:

>CG50817-04-prot 224 aa
MSDEDSCVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGR
QAPEEWSVGLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYA
DHHLPDGERGWVLGRARPGAGISSLQTVPVTLLGPRACSRLHAAPGGDGSPILPGMVCTS
AVGELPSCEANQPAADRGPGHSQEQENAGRQMALLPLSSPPCHV (SEQ ID NO:43)

Figure 17. Nucleotide and Protein Sequences for CG50817-05.

. Nucleotide sequence encoding the Peptidase-like protein of the invention.

>CG50817-05
CGCTGGGCCTCTGTCCTG**ATG**CTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTAC 60
CTGGCCTGGATCCTGTTCTTCGTGCTCTATGATTTCTGCATTGTTTGTATCACCACCTAT 120
GCTATCAACGTGAGCCTGATGTGGCTCAGTTTCCGGAAGGTCCAAGAACCCCAGGGCCAA 180
CCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCCTGGCAGGCCAGTGTGAGG 240
AGGCAAGGAGCCCACATCTGCAGCGGCTCCCTGGTGGCAGACACCTGGGTCCTCACTGCT 300
GCCCACTGCTTTGAAAAGGCAGCAGCAACAGAACTGAATTCCTGCGTGAGGGACTCAGCC 360
CCTGGGGCCGAAGAGGTGGGGGTGGCTGCCCTGCAGTTGCCCAGGGCCTATAACCACTAC 420
AGCCAGGGCTCAGACCTGGCCCTGCTGCAGCTCGCCCACCCCACGACCCACACACCCCTC 480
TGCCTGCCCCAGCCCGCCCATCGCTTCCCCTTTGGAGCCTCCTGCTGGGCCACTGGCTGG 540
GATCAGGACACCAGTGATGCTCCTGGGACCCTACGCAATCTGCGCCTGCGTCTCATCAGT 600
CGCCCCACATGTAACTGTATCTACAACCAGCTGCACCAGCGACACCTGTCCAACCCGGCC 660
CGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTGGGGTGCAGGGCCCCTGTCAGGGAGAT 720
TCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGACACTGGGTTCAGGCTGGCATCATC 780
AGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTGTGCTGCTGACCAACACAGCTGCT 840
CACAGTTCCTGGCTGCAGGCTCGAGTTCAGGGGGCAGCTTTCCTGGCCCAGAGCCCAGAG 900
ACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGATCCTTGAGGACAGCAGGT 960
CCCCAGGCAGGAGCACCCTCCCCATGGCCCTGGGAGGCCAGGCTGATGCACCAGGGACAG 1020
CTGGCCTGTGGCGGAGCCCTGGTGTCAGAGGAGGCGGTGCTAACTGCTGCCCACTGCTTC 1080
ATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGACCGGAGGAGTGG 1140
GGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCTGAGGGGGGCTACGACATG 1200
GCCCTCCTGCTGCTGGCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCCTCTGCCTG 1260
CCCTATGCTGACCACCACCTGCCTGATGGGGAGCGTGGCTGGGTTCTGGGACGGGCCCGC 1320
CCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAGGGCC 1380
TGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGGTG 1440
TGTACCAGTGCTGTGGGTGAGCTGCCCAGCTGTGAGGCCAACCAACCAGCTGCTGACAGG 1500
GGACCTGGCCATTCTCAGGAACAAGAGAATGCAGGCAGGCAAATGGCATTACTGCCCCTG 1560
TCCTCCCCACCCTGTCATGTG**TGA**TTCCAGGC 1592
(SEQ ID NO:44)

Protein sequence encoded by the coding sequence shown above.

>CG50817-05
MLLSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQGQPKPQEG 60
NTVPGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSCVRDSAPGAEEV 120
GVAALQLPRAYNHYSQGSDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTSD 180
APGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSGGPVL 240
CLEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSD 300
EDSCVACGSLRTAGPQAGAPSPWPWEARLMHQGQLACGGALVSEEAVLTAAHCFIGRQAP 360
EEWSVGLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYADHH 420
LPDGERGWVLGRARPGAGISSLQTVPVTLLGPRACSRLHAAPGGDGSPILPGMVCTSAVG 480
ELPSCEANQPAADRGPGHSQEQENAGRQMALLPLSSPPCHV 521
(SEQ ID NO:45)

Figure 18. Nucleotide and Protein Sequences for CG50817-06.

Nucleotide sequence encoding the Peptidase-like protein of the invention.

```
>CG50817-06
AGCGACACCTGTCCAACCCGGCCCGGCCTGGGATGCTATGTGGGGGCCCCCAGCCTGGGG   60
TGCAGGGCCCCTGTCAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGAC   120
ACTGGGTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCCAGGAGGACGCTCCTG   180
TGCTGCTGACCAACACAGCTGCTCACAGTTCCTGGCTGCAGGCTCGAGTTCAGGGGGCAG   240
CTTTCCTGGCCCAGAGCCCAGAGACCCCGGAGATGAGTGATGAGGACAGCTGTGTAGCCT   300
GTGGATCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCCTCCCCATGGCCCTGGGAGG   360
CCAGGCTGATGCACCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTCAGAGGAGGCGG   420
TGCTAACTGCTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGC.  480
TGGGGACCAGACCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCC   540
ACCCTGAGGGGGGCTACGACATGGCCCTCCTGCTGCTGGCCCAGCCTGTGACACTGGGAG   600
CCAGCCTGCGGCCCCTCTGCCTGCCCTATGCTGACCACCACCTGCCTGATGGGGAGCGTG   660
GCTGGGTTCTGGGACGGGCCCGCCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCG   720
TGACCCTCCTGGGGCCTAGGGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCA   780
GCCCTATTCTGCCGGGGATGGTGTGTACCAGTGCTGTGGGTGAGCTGCCCAGCTGTGAGG   840
CCAACCAACCAGCTGCTGACAGGGGACCTGGCCATTCTCAGGAACAAGAGAATGCAGGCA   900
GGCAAATGGCATTACTGCCCCTGTCCTCCCCACCCTGTCATGTGTGATTCCAGGCACCAG   960
GGCAGGCCCAGAAGCCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCAC   1020
TCCCCACCCTGCAGGACAGGGGTGTCTGTGGACACTCCCACACCCAACTCTGCTACCAAG   1080
CAGGCGTCTCAGCTTTCCTCCTCCTTTACCCTTTCAGATACAATCACGCCAGCCACGTTG   1140
TTTTGAAAATTTCTTTTTTTGGGGGGCAGCAGTTTTCCTTTTTTTAAACTTAAATAAATT   1200
(SEQ ID NO:46)
```

Protein sequence encoded by the coding sequence shown above.

```
>CG50817-06
MLCGGPQPGVQGPCQGDSGGPVLCLEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSS   60
WLQARVQGAAFLAQSPETPEMSDEDSCVACGSLRTAGPQAGAPSPWPWEARLMHQGQLAC   120
GGALVSEEAVLTAAHCFIGRQAPEEWSVGLGTRPEEWGLKQLILHGAYTHPEGGYDMALL   180
LLAQPVTLGASLRPLCLPYADHHLPDGERGWVLGRARPGAGISSLQTVPVTLLGPRACSR   240
LHAAPGGDGSPILPGMVCTSAVGELPSCEANQPAADRGPGHSQEQENAGRQMALLPLSSP   300
PCHV                                                           304
(SEQ ID NO:47)
```

Figure 19. Nucleotide and Protein Sequences For CG51099-03.

Nucleotide sequence encoding the Serine Protease-like protein of the invention.

>CG51099-03
CGGAGAGACGCAGTCGGCTGCCACCCCGGG**ATG**GGTCGCTGGTGCCAGACCGTCGCGCGC 60
GGGCAGCGCCCCCGGACGTCTGCCCCCTCCCGCGCCGGTGCCCTGCTGCTGCTGCTTCTG 120
TTGCTGAGGTCTGCAGGTTGCTGGGGCGCAGGGGAAGCCCCGGGGGCGCTGTCCACTGCT 180
GATCCCGCCGACCAGAGCGTCCAGTGTGTCCCCAAGGCCACCTGTCCTTCCAGCCGGCCT 240
CGCCTTCTCTGGCAGACCCCGACCACCCAGACACTGCCCTCGACCACCATGGAGACCCAA 300
TTCCCAGTTTCTGAAGGCAAAGTCGACCCATACCGCTCCTGTGGCTTTTCCTACGAGCAG 360
GACCCCACCCTCAGGGACCCAGAAGCCGTGGCTCGGCGGTGGCCCTGGATGGTCAGCGTG 420
CGGGCCAATGGCACACACATCTGTGCCGGCACCATCATTGCCTCCCAGTGGGTGCTGACT 480
GTGGCCCACTGCCTGATCTGGCGTGATGTTATCTACTCAGTGAGGGTGGGGAGTCCGTGG 540
ATTGACCAGATGACGCAGACCGCCTCCGATGTCCCGGTGCTCCAGGTCATCATGCATAGC 600
AGGTACCGGGCCCAGCGGTTCTGGTCCTGGGTGGGCCAGGCCAACGACATCGGCCTCCTC 660
AAGCTCAAGCAGGAACTCAAGTACAGCAATTACGTGCGGCCCATCTGCCTGCCTGGCACG 720
GACTATGTGTTGAAGGACCATTCCCGCTGCACTGTGACGGGCTGGGGACTTTCCAAGGCT 780
GACGGCATGTGGCCTCAGTTCCGGACCATTCAGGAGAAGGAAGTCATCATCCTGAACAAC 840
AAAGAGTGTGACAATTTCTACCACAACTTCACCAAAATCCCCACTCTGGTTCAGATCATC 900
AAGTCCCAGATGATGTGTGCGGAGGACACCCACAGGGAGAAGTTCTGCTATGAGCTAACT 960
GGAGAGCCCTTGGTCTGCTCCATGGAGGGCACGTGGTACCTGGTGGGATTGGTGAGCTGG 1020
GGTGCAGGCTGCCAGAAGAGCGAGGCCCCACCCATCTACCTACAGGTCTCCTCCTACCAA 1080
CACTGGATCTGGGACTGCCTCAACGGGCAGGCCCTGGCCCTGCCAGCCCCATCCAGGACC 1140
CTGCTCCTGGCACTCCCACTGCCCCTCAGCCTCCTTGCTGCCCTC**TGA**CTCTGTGTGCCC 1200
TCCCTCACTTGTGA 1214
(SEQ ID NO:48)

Protein sequence encoded by the nucleotide sequence shown above.

>CG51099-03
MGRWCQTVARGQRPRTSAPSRAGALLLLLLLLRSAGCWGAGEAPGALSTADPADQSVQCV 60
PKATCPSSRPRLLWQTPTTQTLPSTTMETQFPVSEGKVDPYRSCGFSYEQDPTLRDPEAV 120
ARRWPWMVSVRANGTHICAGTIIASQWVLTVAHCLIWRDVIYSVRVGSPWIDQMTQTASD 180
VPVLQVIMHSRYRAQRFWSWVGQANDIGLLKLKQELKYSNYVRPICLPGTDYVLKDHSRC 240
TVTGWGLSKADGMWPQFRTIQEKEVIILNNKECDNFYHNFTKIPTLVQIIKSQMMCAEDT 300
HREKFCYELTGEPLVCSMEGTWYLVGLVSWGAGCQKSEAPPIYLQVSSYQHWIWDCLNGQ 360
ALALPAPSRTLLLALPLPLSLLAAL 385 (SEQ ID NO:49)

Figure 20. Nucleotide and Protein Sequences For CG57051-04.

Nucleotide sequence encoding the Angiopoietin-like protein, CG57051-04.

```
>CG57051-04
TGCGGATCCTCACACGACTGTGATCCGATTCTTTCCAGCGGCTTCTGCAACCAAGCGGGT  60
CTTACCCCCGGTCCTCCGCGTCTCCAGTCCTCGCACCTGGAACCCCAACGTCCCCGAGAG  120
TCCCCGAATCCCCGCTCCCAGGCTACCTAAGAGGATGAGCGGTGCTCCGACGGCCGGGGC  180
AGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCTAGATCTGGACCCGTGCA  240
GTCCAAGTCGCCGCGCTTTGCGTCCTGGGACGAGATGAATGTCCTGGCGCACGGACTCCT  300
GCAGCTCGGCCAGGGGCTGCGCGAACACGCGGAGCGCACCCGCAGTCAGCTGAGCGCGCT  360
GGAGCGGCGCCTGAGCGCGTGCGGGTCCGCCTGTCAGGGAACCGAGGGGTCCACCGACCT  420
CCCGTTAGCCCCTGAGAGCCGGGTGGACCCTGAGGTCCTTCACAGCCTGCAGACACAACT  480
CAAGGCTCAGAACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCCAGCAGCAGCGGCA  540
CCTGGAGAAGCAGCACCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCCTGGACCA  600
CAAGCACCTAGACCATGAGGTGGCCAAGCCTGCCCGAAGAAAGAGGCTGCCCGAGATGGC  660
CCAGCCAGTTGACCCGGCTCACAATGTCAGCCGCCTGCACCGAGGCTGGTGGTTTGGCAC  720
CTGCAGCCATTCCAACCTCAACGGCCAGTACTTCCGCTCCATCCCACAGCAGCGGCAGAA  780
GCTTAAGAAGGGAATCTTCTGGAAGACCTGGCGGGGCCGCTACTACCCGCTGCAGGCCAC  840
CACCATGTTGATCCAGCCCATGGCAGCAGAGGCAGCCTCCTAGCGTCCTGGCTGGGCCTG  900
GTCCCAGGCCCACGAAAGACGGTGACTCTTGGCTCTG  937  (SEQ ID NO:50)
```

Protein sequence encoded by the nucleotide sequence shown above.

```
>CG57051-04
MSGAPTAGAALMLCAATAVLLSARSGPVQSKSPRFASWDEMNVLAHGLLQLGQGLREHAE  60
RTRSQLSALERRLSACGSACQGTEGSTDLPLAPESRVDPEVLHSLQTQLKAQNSRIQQLF  120
HKVAQQQRHLEKQHLRIQHLQSQFGLLDHKHLDHEVAKPARRKRLPEMAQPVDPAHNVSR  180
LHRGWWFGTCSHSNLNGQYFRSIPQQRQKLKKGIFWKTWRGRYYPLQATTMLIQPMAAEA  240
AS  242  (SEQ ID NO:51)
```

Figure 21. Nucleotide and Protein Sequences For CG57051-05.

Nucleotide sequence encoding the Angiopoietin-like protein, CG57051-05.

>CG57051-05
CTTCGTCTCCAGTCCTCGCACCTGGAACCCCAACGTCCCCGAGAGTCCCCGAATCCCCGC 60
TCCCAGGCTACCTAAGAGG**ATG**AGCGGCGCTCCGACGGCCGGGGCAGCCCTGATGCTCTG 120
CGCCGCCACCGCCGTGCTACTGAGCGCTCAGGGCGGACCCGTGCAGTCCAAGTCGCCGCG 180
CTTTGCGTCCTGGGACGAGATGAATGTCCTGGCGCACGGACTCCTGCAGCTCGGCCAGGG 240
GCTGCGCGAACACGCGGAGCGCACCCGCAGTCAGCTGAGCGCGCTGGAGCGGCGCCTGAG 300
CGCGTGCGGGTCCGCCTGTCAGGGAACCGAGGGGTCCACCGACCTCCCGTTAGCCCCTGA 360
GAGCCGGGTGGACCCTGAGGTCCTTCACAGCCTGCAGACACAACTCAAGGCTCAGAACAG 420
CAGGATCCAGCAACTCTTCCACAAGGTGGCCCAGCAGCAGCGGCACCTGGAGAAGCAGCA 480
CCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCCTGGACCACAAGCACCTAGACCA 540
TGAGGGTGGCAAGCCTGCCCGAAGAAAGAGGCTGCCCGAGATGGCCCAGCCAGTTGACCC 600
GGCTCACAATGTCAGCCGCCTGCACCATGGAGGCTGGACAGTAATTCAGAGGCGCCACGA 660
TGGCTCAGTGGACTTCAACCGGCCCTGGGAAGCCTACAAGGCGGGGTTTGGGGATCCCCA 720
CGGCGAGTTCTGGCTGGGTCTGGAGAAGGTGCATAGCATCATGGGGGACCGCAACAGCCG 780
CCTGGCCGTGCAGCTGCGGGACTGGGATGGCAACGCCGAGTTGCTGCAGTTCTCCGTGCA 840
CCTGGGTGGCGAGGACACGGCCTATAGCCTGCAGCTCACTGCACCCGTGGCCGGCCAGCT 900
GGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGTACCCTTCTCCACTTGGGACCAGGA 960
TCACGACCTCCGCAGGGACAAGAACTGCGCCAAGAGCCTCTCTGGAGGCTGGTGGTTTGG 1020
CACCTGCAGCCATTCCAACCTCAACGGCCAGTACTTCCGCTCCATCCCACAGCAGCGGCA 1080
GAAGCTTAAGAAGGGAATCTTCTGGAAGACCTGGCGGGGCCGCTACTACCCGCTGCAGGC 1140
CACCACCATGTTGATCCAGCCCATGGCAGCAGAGGCAGCCTCC**TAG**CGTCCTGGCTGGGC 1200
CTGGTCCCAGGCCCACGAAAGAGGTGACTCTTGGCTCTG 1239 (SEQ ID NO:52)

Protein sequence for Angiopoietin-like protein, CG57051-05.

>CG57051-05
MSGAPTAGAALMLCAATAVLLSAQGGPVQSKSPRFASWDEMNVLAHGLLQLGQGLREHAE 60
RTRSQLSALERRLSACGSACQGTEGSTDLPLAPESRVDPEVLHSLQTQLKAQNSRIQQLF 120
HKVAQQQRHLEKQHLRIQHLQSQFGLLDHKHLDHEGGKPARRKRLPEMAQPVDPAHNVSR 180
LHHGGWTVIQRRHDGSVDFNRPWEAYKAGFGDPHGEFWLGLEKVHSIMGDRNSRLAVQLR 240
DWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPPSGLSVPFSTWDQDHDLRRD 300
KNCAKSLSGGWWFGTCSHSNLNGQYFRSIPQQRQKLKKGIFWKTWRGRYYPLQATTMLIQ 360
PMAAEAAS 368 (SEQ ID NO:53)

Figure 22. Nucleotide and Protein Sequences For CG57051-02.

Nucleotide sequence encoding the Angiopoietin-like protein of the invention.

```
>CG57051_02
TGCGGATCCTCACACGACTGTGATCCGATTCTTTCCAGCGGCTTCTGCAACCAAGCGGGT      60
CTTACCCCCGGTCCTCCGCGTCTCCAGTCCTCGCACCTGGAACCCCAACGTCCCCGAGAG     120
TCCCCGAATCCCCGCTCCCAGGCTACCTAAGAGGATGAGCGGTGCTCCGACGGCCGGGGC     180
AGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCTAGATCTGGACCCGTGCA     240
GTCCAAGTCGCCGCGCTTTGCGTCCTGGGACGAGATGAATGTCCTGGCGCACGGACTCCT     300
GCAGCTCGGCCAGGGGCTGCGCGAACACGCGGAGCGCACCCGCAGTCAGCTGAGCGCGCT     360
GGAGCGGCGCCTGAGCGCGTGCGGGTCCGCCTGTCAGGGAACCGAGGGGTCCACCGACCT     420
CCCGTTAGCCCCTGAGAGCCGGGTGGACCCTGAGGTCCTTCACAGCCTGCAGACACAACT     480
CAAGGCTCAGAACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCCAGCAGCAGCGGCA     540
CCTGGAGAAGCAGCACCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGCCTCCTGGACCA     600
CAAGCACCTAGACCATGAGGTGGCCAAACCTGCCCGAAGAAAGAGGCTGCCCGAGATGGC     660
CCAGCCAGTTGACCCGGCTCACAATGTCAGCCGCCTGCACCATGGAGGCTGGACAGTAAT     720
TCAGAGGCGCCACGATGGCTCAATGGACTTCAACCGGCCCTGGGAAGCCTACAAGGCGGG     780
GTTTGGGGATCCCCACGGCGAGTTCTGGCTGGGTCTGGAGAAGGTGCATAGCATCACGGG     840
GGACCGCAACAGCCGCCTGGCCGTGCAGCTGCGGGACTGGGATGGCAACGCCGAGTTGCT     900
GCAGTTCTCCGTGCACCTGGGTGGCGAGGACACGGCCTATAGCCTGCAGCTCACTGCACC     960
CGTGGCCGGCCAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGTACCCTTCTC    1020
CACTTGGGACCAGGATCACGACCTCCGCAGGGACAAGAACTGCGCCAAGAGCCTCTCTGC    1080
CCCATCGGTGGCTCAAAGACCTGACCATGTTCCCTCTCCCCTGACCCCGGCAGGAGGCTG    1140
GTGGTTTGGCACCTGCAGCCATTCCAACCTCAACGGCCAGTACTTCCGCTCCATCCCACA    1200
GCAGCGGCAGAAGCTTAAGAAGGGAATCTTCTGGAAGACCTGGCGGGGCCGCTACTACCC    1260
GCTGCAGGCCACCACCATGTTGATCCAGCCCATGGCAGCAGAGGCAGCCTCCTAG   1315
(SEQ ID NO:54)
```

Protein sequence for CG57051-02.

```
>CG57051_02
MSGAPTAGAALMLCAATAVLLSARSGPVQSKSPRFASWDEMNVLAHGLLQLGQGLREHAE      60
RTRSQLSALERRLSACGSACQGTEGSTDLPLAPESRVDPEVLHSLQTQLKAQNSRIQQLF     120
HKVAQQQRHLEKQHLRIQHLQSQFGLLDHKHLDHEVAKPARRKRLPEMAQPVDPAHNVSR     180
LHHGGWTVIQRRHDGSMDFNRPWEAYKAGFGDPHGEFWLGLEKVHSITGDRNSRLAVQLR     240
DWDGNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPPSGLSVPFSTWDQDHDLRRD     300
KNCAKSLSAPSVAQRPDHVPSPLTPAGGWWFGTCSHSNLNGQYFRSIPQQRQKLKKGIFW     360
KTWRGRYYPLQATTMLIQPMAAEAAS    386   (SEQ ID NO:55)
```

Figure 23. Nucleotide and Protein Sequences For CG57051-03.

Nucleotide sequence encoding the Angiopoietin-like protein, CG57051-03.

```
>CG57051-03
CCCCGAGAGTCCCCGAATCCCCGCTCCCAGGCTACCTAAGAGGATGAGCGGTGCTCCGAC     60
GGCCGGGGCAGCCCTGATGCTCTGCGCCGCCACCGCCGTGCTACTGAGCGCTCAGGGCGG    120
ACCCGTGCAGTCCAAGTCGCCGCGCTTTGCGTCCTGGGACGAGATGAATGTCCTGGCGCA    180
CGGACTCCTGCAGCTCGGCCAGGGGCTGCGCGAACACGCGGAGCGCACCCGCAGTCAGCT    240
GAGCGCGCTGGAGCGGCGCCTGAGCGCGTGCGGGTCCGCCTGTCAGGGAACCGAGGGGTC    300
CACCGACCTCCCGTTAGCCCCTGAGAGCCGGGTGGACCCTGAGGTCCTTCACAGCCTGCA    360
GACACAACTCAAGGCTCAGAACAGCAGGATCCAGCAACTCTTCCACAAGGTGGCCCAGCA    420
GCAGCGGCACCTGGAGAAGCAGCACCTGCGAATTCAGCATCTGCAAAGCCAGTTTGGCCT    480
CCTGGACCACAAGCACCTAGACCATGAGGTGGCCAAGCCTGCCCGAAGAAAGAGGCTGCC    540
CGAGATGGCCCAGCCAGTTGACCCGGCTCACAATGTCAGCCGCCTGCACCATGGAGGCTG    600
GACAGTAATTCAGAGGCGCCACGATGGCTCAGTGGACTTCAACCGGCCCTGGGAAGCCTA    660
CAAGGCGGGGTTTGGGGATCCCCACGGCGAGTTCTGGCTGGGTCTGGAGAAGGTCCATAG    720
CATCACGGGGGACCGCAACAGCCGCCTGGCCGTGCAGCTGCGGGACTGGGATGACAACGC    780
CGAGTTGCTGCAGTTCTCCGTGCACCTGGGTGGCGAGGACACGGCCTATAGCCTGCAGCT    840
CACTGCACCCGTGGCCGGCCAGCTGGGCGCCACCACCGTCCCACCCAGCGGCCTCTCCGT    900
ACCCTTCCCCACTTGGGACCAGGATCACGACCTCCGCAGGGACAAGAACTGCGCCAAGAG    960
CCTCTCTGGAGGCTGGTGGTTTGGCACCTGCAGCCATTCCAACCTCAACGGCCAGTACTT   1020
CCGCTCCATCCCACAGCAGCGGCAGAAGCTTAAGAAGGGAATCTTCTGGAAGACCTGGCG   1080
GGGCCGCTACTACCCGCTGCAGGCCACCACCATGTTGATCCAGCCCATGGCAGCAGAGGC   1140
AGCCTCCTAG    1150 (SEQ ID NO:56)
```

Protein sequence for CG57051-03.

```
>CG57051-03
MSGAPTAGAALMLCAATAVLLSAQGGPVQSKSPRFASWDEMNVLAHGLLQLGQGLREHAE     60
RTRSQLSALERRLSACGSACQGTEGSTDLPLAPESRVDPEVLHSLQTQLKAQNSRIQQLF    120
HKVAQQQRHLEKQHLRIQHLQSQFGLLDHKHLDHEVAKPARRKRLPEMAQPVDPAHNVSR    180
LHHGGWTVIQRRHDGSVDFNRPWEAYKAGFGDPHGEFWLGLEKVHSITGDRNSRLAVQLR    240
DWDDNAELLQFSVHLGGEDTAYSLQLTAPVAGQLGATTVPPSGLSVPFPTWDQDHDLRRD    300
KNCAKSLSGGWWFGTCSHSNLNGQYFRSIPQQRQKLKKGIFWKTWRGRYYPLQATTMLIQ    360
PMAAEAAS    368 (SEQ ID NO:57)
```